

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2005, 11:40:32 ; Search time 8758 Seconds

(without alignments)
12099.900 Million cell updates/sec

Title: US-09-068-751-1

Perfect score: 2784

Sequence: 1 gaagtgaatctgaataatt.....cagttgcgcgcgcgaaca 2784

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gss1:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	398.2	14.3	414	2	AW936842 PM4-DT002
C 2	291	10.5	548	2	BE110694 UT-R-BJ1-
C 3	269.8	9.7	855	8	BZ756044 PUBBL24TB
C 4	269.8	9.7	855	8	BZ756044 PUBBL24TB
5	267.8	9.6	949	8	BZ677942 PUBFUS5TD
6	256.2	9.2	867	8	BZ681727 PUBH42TD
7	253.8	9.1	343	6	BZ683904 PUBH053TD
8	253.8	9.1	343	6	CB693457 AMGNNUC:N
9	253.8	9.1	574	6	CB608609 AMGNNUC:N
10	253.8	9.1	659	6	CB579321 AMGNNUC:N
11	252.4	9.1	973	8	CB546382 AMGNNUC:N
C 12	244.6	8.8	289	8	BZ701200 PUBMP50TD
C 13	241.6	8.7	697	6	BZ670825 PUBH64TD
14	240.4	8.6	841	8	BY752869 BY752869
15	236.4	8.5	514	8	BZ669574 PUBC087TD
16	235.8	8.5	856	8	BZ089457 CH230-1C1
17	235.6	8.5	747	9	BZ682015 PUBJF02TD
18	231.8	8.3	687	8	AG548014 Mus muscu
19	230.4	8.3	578	6	BZ683170 PUBC064TD
20	226.8	8.1	520	8	CA752540 UI-M-FOO-
C 21	226.6	8.1	3561	3	BZ128636 CH230-440
C 22	225.4	8.1	854	9	AK030040 Mus muscu
C 23	224.8	8.1	541	6	CB252970 Forward s
C 24	220.8	7.9	492	8	CB612562 AMGNNUC:N
C 25	220.8	7.9	492	8	BZ684949 PUBCH20TD

25	220.8	7.9	902	8	BZ670418
26	218	7.8	669	8	AZ392101
C 27	217	7.8	879	9	CNS03151
28	212.8	7.6	796	8	BZ669209
29	212.2	7.6	643	8	BZ207577
C 30	208.8	7.5	523	8	AZ383375
31	208.8	7.5	1197	9	AG346705
32	208.6	7.5	739	9	AG550880
33	206.2	7.4	750	9	AG601211
34	205.4	7.4	997	9	CNS005TE
C 35	204	7.3	571	8	BZ182096
36	202.4	7.3	533	8	BZ173741
C 37	202.2	7.3	567	8	AZ895651
38	202	7.3	776	5	BUS58894
C 39	199.4	7.2	776	9	CR133916
40	199.4	7.2	815	9	CNS020OC
41	198	7.1	732	9	AG404322
42	197.6	7.1	726	9	AG402131
43	193.2	6.9	776	9	AG561854
44	192.6	6.9	650	8	AZ337339
45	190.6	6.8	783	9	AG573654

ALIGNMENTS

RESULT 1
LOCUS AW936842 414 bp mRNA linear EST 30-MAY-2000
DEFINITION PM4-DT0025-031299-001-B09 DT0025 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW936842
VERSION AW936842.1 GI:8112256
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 414)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,U.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tit=at2=PM4-DT0025-031299-001-B09&ts=1999-12-03&ts=1)
Seq primer: puc 18 forward
High quality sequence start: 34
High quality sequence stop: 406.

FEATURES
source
1..414
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="DT0025"
/note="Organ: denis drash; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products

derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 14.3%; Score 398.2; DB 2; Length 414;
Best Local Similarity 98.1%; Pred. No. 5e-74; 8; Indels 0; Gaps 0;
Matches 403; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 228 AAATCTCACACGACAAAGATTGGAGTTCACAGGATGGGTGTCGCCAAGACCTTAGGGA 287
Db |||||
2 AAATATCACGGGCTTAAGAGTTTGGAGTTCACAGGATGGGTGTCGCCAAGACCTTAGGGA 61
QY 288 CAGNAGTTGTGAGCCCTGTCGCCGACCCCTGTGGACTGTGAGAGACGACAGTCCAC 347
Db |||||
62 CAGNAGTTGTGAGCCCTGTCGCCGACCCCTGTGGACTGTGAGAGACGACAGTCCAC 121
QY 348 CCCCAGGCTTCTTAGACCCACCCCGGTTTCCAGCATCTTCTGCGAGACCGGACC 407
Db |||||
122 CCCCAGGCTTCTTAGACCCACCCCGGTTTCCAGCATCTTCTGCGAGACCGGACC 181
QY 408 CTGGCTGAAAGTACAGAAACCTAGAGTCTGACGCCCATGTGGTGGCCGCCCATGTTT 467
Db |||||
182 CTGGCTGAAAGTACAGAAACCTAGAGTCTGACGCCCATGTGGTGGCCGCCCATGTTT 241
QY 468 CCAGNATCTCTGGTCTAGGATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGACAG 527
Db |||||
242 CCAGNATCTCTGGTCTAGGATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGACAG 301
QY 528 TTAGCATGTTCTATGTCTACTGCGAGGAGCAGGACCAACAGAGGTCATGAAGATCCACAG 587
Db |||||
302 TTAGCATGTTCTATGTCTACTGCGAGGAGCAGGACCAACAGAGGTCATGAAGATCCACAG 361
QY 588 GGGCTCCGGTCCGAGGCCCTTGGGTTTATCACCAGATGTTTCCACCCA 638
Db |||||
362 GGGCTCCGGTCCGAGGCCCTTGGGTTTATCACCAGATGTTTCCACCCA 412

RESULT 2

BE110694/c
LOCUS BZ756044/c
DEFINITION PUBBL24TB ZM_0.6_1.0_KB zea mays genomic clone ZMBE149C24,
UI-R-BJ1-avu-h-11-0-UI sl UI-R-BJ1 Rattus norvegicus cDNA clone
UI-R-BJ1-avu-h-11-0-UI 3', mRNA sequence.
BE110694
ACCESSION BZ756044.1 GI:28908425
VERSION BZ756044.1
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 548)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
MEDLINE 97044477
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRP, Iowa City, IA 52242, USA
Tel: 319 335 8250
Fax: 319 335 9565
Email: bento-soares@uiowa.edu
The sequence contained an oligo-dT track that was present in the oligonucleotide that was used to prime the synthesis of first strand cDNA and therefore this may represent a bonafide poly A tail. The sequence tag present in the cDNA between the NotI site and the oligo-dT track served to identify it as a clone from the normalized AV canal at 15 dpc library cDNA Library Preparation:

FEATURES
source

1..548
/organism="Rattus norvegicus"
/mol_type="mRNA"
/strain="Sprague-Dawley"
/db_xref="taxon:10116"
/clone="UI-R-BJ1-avu-h-11-0-UI"
/lab_host="DHI0B (Life Technologies)"
/clone_lib="UI-R-BJ1"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; The UI-R-BJ1 library is a subtracted library derived from the following tissues: atrium at 16.5 dpc, ventricle at 16.5 dpc, AV canal at 16.5 dpc, atrium at 15 dpc, ventricle at 15 dpc, AV canal at 15 dpc, ventricle at 13 dpc, and adult heart. For a detailed description of the library from which this clone was derived, please visit our web site at ratetst.eng.uiowa.edu. The subtraction has been previously described in (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG_TISSUE=AV canal at 15 dpc
TAG_LIB=UI-R-BJ1
TAG_SEQ=GAAGG"

ORIGIN

Query Match 10.5%; Score 291; DB 2; Length 548;
Best Local Similarity 100.0%; Pred. No. 3.5e-51;
Matches 291; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2132 TTTTCCCCATTAGACAATGGCAGGACCCAGACACAGAGCATCGTTCCAGGCCAGGCC 2191
Db |||||
538 TTTTCCCCATTAGACAATGGCAGGACCCAGACACAGAGCATCGTTCCAGGCCAGGCC 479
QY 2192 CAGCCACATGCTCTTTAAACCTTTGAAGGCATTTTTCGGTCTCACGTTGTCACCCAGCGGG 2251
Db |||||
478 CAGCCACATGCTCTTTAAACCTTTGAAGGCATTTTTCGGTCTCACGTTGTCACCCAGCGGG 419
QY 2252 TGTCCGACTTTGAACGGCTCTTACTTTCAGAAAGACCGCATGGGGTGGGGGCTTAGGTG 2311
Db |||||
418 TGTCCGACTTTGAACGGCTCTTACTTTCAGAAAGACCGCATGGGGTGGGGGCTTAGGTG 359
QY 2312 GCCTCTGCCTCACCTACAACTGCCAAAGTGGTCAATGGGTTATTTTAAACCCAGGGNA 2371
Db |||||
358 GCCTCTGCCTCACCTACAACTGCCAAAGTGGTCAATGGGTTATTTTAAACCCAGGGNA 299
QY 2372 GAGGTATTTATTTGTTTCCACAGCGGGCGGCCAGCAGGCTCCTTGAATTC 2422
Db |||||
298 GAGGTATTTATTTGTTTCCACAGCGGGCGGCCAGCAGGCTCCTTGAATTC 248

RESULT 3

BE110694/c
LOCUS BZ756044/c
DEFINITION PUBBL24TB ZM_0.6_1.0_KB zea mays genomic clone ZMBE149C24,
genomic survey sequence.
ACCESSION BZ756044
VERSION BZ756044.1 GI:28908425
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
REFERENCE 1 (bases 1 to 855)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennettzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)

BZ756044 855 bp DNA linear GSS 10-MAR-2003
PUBBL24TB ZM_0.6_1.0_KB zea mays genomic clone ZMBE149C24,
genomic survey sequence.
BZ756044
BZ756044.1 GI:28908425
GSS.
Zea mays
Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Zea.
1 (bases 1 to 855)
Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T., Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and Bennettzen,J.
Maize Genomics Consortium
Unpublished (2003)

COMMENT

Other_GSSs: PUBBL24TD
 Contact: Cathy Whitelaw
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TR
 Class: sheared ends.

FEATURES

source
 1. .855
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMBTa149C24"
 /clone_lib="ZM_0.6_1.0_KB"
 /note="Vector: pCR4-TOFO; Site 1: EcoRI; 0.6-1.0 kb high
 CoT selected genomic DNA library"

ORIGIN

Query Match 9.7%; Score 269.8; DB 8; Length 855;
 Best Local Similarity 94.0%; Pred. No. 1.2e-46;
 Matches 313; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
 QY 2443 TCCGGTACTGTTGGTAAATGGAAGACGCCAAACATAAAGAAAGCCCGCGCCATTTC 2502
 Db 565 TCCGGTACTGTTGGTAAATGGAAGACGCCAAACATAAAGAAAGCCCGCGCCATTTC 506
 QY 2503 TATCCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAGAGATAGAGATACGCC 2562
 Db 505 TATCCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAGAGATAGAGATACGCC 446
 QY 2563 CTGGTTCCTGGAACATTTGCTTTTACAGATGCACATATCGAGTGAACATCAGTTTCGCG 2622
 Db 445 CTGGTTCCTGGAACATTTGCTTTTACAGATGCACATATCGAGTGAACATCAGTTTCGCG 386
 QY 2623 GAATACACTCGAATGTCGTTTCGTTGGCAGAGCTATGAACAGATATGGCTGAATAC 2682
 Db 385 GAATACACTCGAATGTCGTTTCGTTGGCAGAGCTATGAACAGATATGGCTGAATAC 327
 QY 2683 AAATCAGAGATCGTGTATGAGTGAACATCTCTTCAATTTCTTTATGCCGGTGTGG 2742
 Db 326 AAATCAGAGATCGTGTATGAGTGAACATCTCTTCAATTTCTTTATGCCGGTGTGG 268
 QY 2743 GCCCGTTATTTATCCGAGTTGCGAGTTGCGGCC 2775
 Db 267 GCCTGTTATTTAT-CCGAGTTGAGTTGCGCCC 236

RESULT 4

BZ677942
 LOCUS PUBFUS5TD ZM_0.6_1.0_KB Zea mays genomic clone ZMBTa045J13,
 DEFINITION genomic survey sequence.
 ACCESSION BZ677942
 VERSION BZ677942.1 GI:28230797
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 892)

REFERENCE
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.

TITLE Maize Genomics Consortium
 JOURNAL Unpublished (2003)
 COMMENT Contact: Cathy Whitelaw

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843

FEATURES

Location/Qualifiers

Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

FEATURES

source
 1. .892
 Location/Qualifiers
 /organism="Zea mays"
 /mol_type="genomic DNA"
 /strain="B73"
 /db_xref="taxon:4577"
 /clone_lib="ZMBTa045J13"
 /clone_lib="ZM_0.6_1.0_KB"
 /note="Vector: pCR4-TOFO; Site 1: EcoRI; 0.6-1.0 kb high
 CoT selected genomic DNA library"

ORIGIN

Query Match 9.7%; Score 269.8; DB 8; Length 892;
 Best Local Similarity 94.0%; Pred. No. 1.2e-46;
 Matches 313; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
 QY 2443 TCCGGTACTGTTGGTAAATGGAAGACGCCAAACATAAAGAAAGCCCGCGCCATTTC 2502
 Db 16 TCCGGTACTGTTGGTAAATGGAAGACGCCAAACATAAAGAAAGCCCGCGCCATTTC 75
 QY 2503 TATCCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAGAGCTAGAGATACGCC 2562
 Db 76 TATCCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAGAGCTAGAGATACGCC 135
 QY 2563 CTGGTTCCTGGAACATTTGCTTTTACAGATGCACATATCGAGTGAACATCAGTTTCGCG 2622
 Db 136 CTGGTTCCTGGAACATTTGCTTTTACAGATGCACATATCGAGTGAACATCAGTTTCGCG 195
 QY 2623 GAATACACTCGAATGTCGTTTCGTTGGCAGAGCTATGAACAGATATGGCTGAATAC 2682
 Db 196 GAATACACTCGAATGTCGTTTCGTTGGCAGAGCTATGAACAGATATGGCTGAATAC 254
 QY 2683 AAATCAGAGATCGTGTATGAGTGAACATCTCTTCAATTTCTTTATGCCGGTGTGG 2742
 Db 255 AAATCAGAGATCGTGTATGAGTGAACATCTCTTCAATTTCTTTATGCCGGTGTGG 313
 QY 2743 GCCCGTTATTTATCCGAGTTGCGAGTTGCGGCC 2775
 Db 314 GCCTGTTATTTAT-CCGAGTTGAGTTGCGCCC 345

RESULT 5

BZ681727
 LOCUS PUBH142TD ZM_0.6_1.0_KB Zea mays genomic clone ZMBTa055H11,
 DEFINITION genomic survey sequence.

ACCESSION BZ681727
 VERSION BZ681727.1 GI:28237508
 KEYWORDS GSS.
 SOURCE Zea mays
 ORGANISM Zea mays

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
 clade; Panicoideae; Andropogoneae; Zea.
 1 (bases 1 to 949)

REFERENCE
 AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
 Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
 Bennetzen,J.

TITLE Maize Genomics Consortium
 JOURNAL Unpublished (2003)
 COMMENT Contact: Cathy Whitelaw

TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA
 Tel: 301-838-5843
 Fax: 301-838-0208
 Email: whitelaw@tigr.org
 Seq primer: TF
 Class: sheared ends.

```

source
1. .949
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBTa055H11"
/notes_lib="ZM_0.6_1.0_KB"
/note_vector="PCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
Cor selected genomic DNA library"

ORIGIN

Query Match          9.6%; Score 267.8; DB 8; Length 949;
Best Local Similarity 94.0%; Pred.No.3.3e-46;
Matches 311; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY   2445  CGGTACTGTTGGTAAATAATGAAGACGCCAATAAAGAAAGCGCCGCCCATCTTA 2504
      |||||
DB    1    CGGTACTGTTGGTAAATAATGAAGACGCCAATAAAGAAAGCGCCGCCCATCTTA 60

QY   2505  TCCTCTAGAGATGGAAACCGCTGGAGAGCCAACTGCATAAGGCTAGAAGAGATACGCCT 2564
      |||||
DB    61   TCCTCTAGAGATGGAAACCGCTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCT 120

QY   2565  GGTTCTCTGAACAATTCGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGGCG 2624
      |||||
DB    121  GGTTCTCTGAACAATTCGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGGCG 180

QY   2625  ATACTTTCGAATCTCGTTTCGGTTGCGAGNAGCTATGAACGATATGGCTGNAATACAA 2684
      |||||
DB    181  ATACTTTCGAATCTCGG-TTCGGTTGCGAAGCTATGAACGATATGGGCTGNAATACAA 239

QY   2685  ATCAGAAATCGTCGTATGTCAGTGAAAATCTCTTTCAAATTCTTTATGCGGTTGTTGGC 2744
      |||||
DB    240  ATCAGAAATCGTCGTATGTCAGTGAAAATCTCT-TTCAAATTCTTTATGCGGTTGTTGGC 298

QY   2745  CGGTATTATTCGGAGTTTCAGTTGCGGCC 2775
      |||||
DB    299  GCGTTATTAT-CGGAGTTTCAGTTGCGGCC 328


RESULT 6
EZ683904
LOCUS
DEFINITION
  BZ683904 Zea mays genomic clone ZMMBTa057I09,
  genomic survey sequence.
ACCESSION
  BZ683904.1 GI:28241625
VERSION
  GSS.
KEYWORDS
  Zea mays
SOURCE
  Zea mays
ORGANISM
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACAD
  clade; Panicoideae; Andropogoneae; Zea.
REFERENCE
  1 (bases 1 to 867)
AUTHORS
  Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
  Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
  Bennetzen,J.
TITLE
  Maize Genomics Consortium
JOURNAL
  Unpublished (2003)
COMMENT
  Contact: Cathy Whitelaw
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA
  Tel: 301-838-5843
  Fax: 301-838-0208
  Email: whitelaw@tigr.org
  Seq primer: TP
  Class: sheared ends.
  Location/Qualifiers
    1..867
      /organism="Zea mays"
      /mol_type="genomic DNA"
      /strain="B73"
      /db_xref="taxon:4577"

```

```
QY 2459 AAATGGAGCGCAAAACATATAAGAGCGCGCGCATTTCTATCTCTAGAGATG 2518
Db 5 AAATGGAAGCGCAAAACATATAAGAGCGCGCGCATTTCTATCTCTAGAGATG 64
QY 2519 GAACCGCTGGAGAGCAACTGCATAAGGCTAGAGAGATACGCCCTGGTTCTTGGAAACAA 2578
Db 65 GAACCGCTGGAGAGCAACTGCATAAGGCTAGAGAGATACGCCCTGGTTCTTGGAAACAA 124
QY 2579 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATGT 2638
Db 125 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATGT 184
QY 2639 CCGTTTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 2698
Db 185 CCG-TTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 243
QY 2699 GTATGCAGTGAACAACTCTCTTTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTATCCG 2758
Db 244 GTATGCAGTGAACAACTCTC-TTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTAT-CG 301
QY 2759 GAGTTGCAGTTGCGGCC 2775
Db 302 GAGTTGCAGTTGCGGCC 318

RESULT 8
CB579321
LOCUS CB579321
DEFINITION CB579321
ACCESSION CB579321
VERSION CB579321
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 574)
AUTHORS Angen EST Program
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00131 row: e column: 2.

FEATURES
source
1..574
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrh4-00131-e2"
/clone_lib="W Rat hypothalamus (10464)"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3 kb fraction 6 and 7"

ORIGIN
Query Match 9.1%; Score 253.8; DB 6; Length 574;
Best Local Similarity 93.7%; Pred. No. 3e-43;
Matches 297; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2459 AAATGGAAGCGCAAAACATATAAGAGCGCGCGCATTTCTATCTCTAGAGATG 2518
Db 5 AAATGGAAGCGCAAAACATATAAGAGCGCGCGCATTTCTATCTCTAGAGATG 64
QY 2519 GAACCGCTGGAGAGCAACTGCATAAGGCTAGAGAGATACGCCCTGGTTCTTGGAAACAA 2578
Db 65 GAACCGCTGGAGAGCAACTGCATAAGGCTAGAGAGATACGCCCTGGTTCTTGGAAACAA 124
QY 2579 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATGT 2638
Db 125 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATGT 184
QY 2639 CCGTTTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 2698
Db 185 CCG-TTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 243
QY 2699 GTATGCAGTGAACAACTCTCTTTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTATCCG 2758
Db 244 GTATGCAGTGAACAACTCTC-TTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTAT-CG 301
```

```
Db 125 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTACGCGGAATACTTCGAAATGT 184
QY 2639 CCGTTTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 2698
Db 185 CCG-TTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 243
QY 2699 GTATGCAGTGAACAACTCTCTTTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTATCCG 2758
Db 244 GTATGCAGTGAACAACTCTC-TTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTAT-CG 301
QY 2759 GAGTTGCAGTTGCGGCC 2775
Db 302 GAGTTGCAGTTGCGGCC 318

RESULT 9
CB579321
LOCUS CB579321
DEFINITION CB579321
ACCESSION CB579321
VERSION CB579321
KEYWORDS EST.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus

REFERENCE 1 (bases 1 to 618)
AUTHORS Angen EST Program
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Angen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00182 row: f column: 1.

FEATURES
source
1..618
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrh4-00182-f1"
/clone_lib="W Rat hypothalamus (10464)"
/notes="Vector: pSPORT1; Site 1: SalI; Site 2: NotI; W Rat hypothalamus adult female Wistar rat avg. insert size 2.3 kb fraction 6 and 7"

ORIGIN
Query Match 9.1%; Score 253.8; DB 6; Length 618;
Best Local Similarity 93.7%; Pred. No. 3e-43;
Matches 297; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2459 AAATGGAAGCGCAAAACATATAAGAGCGCGCGCATTTCTATCTCTAGAGATG 2518
Db 5 AAATGGAAGCGCAAAACATATAAGAGCGCGCGCATTTCTATCTCTAGAGATG 64
QY 2519 GAACCGCTGGAGAGCAACTGCATAAGGCTAGAGAGATACGCCCTGGTTCTTGGAAACAA 2578
Db 65 GAACCGCTGGAGAGCAACTGCATAAGGCTAGAGAGATACGCCCTGGTTCTTGGAAACAA 124
QY 2579 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATGT 2638
Db 125 TTGCTTTTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATGT 184
QY 2639 CCGTTTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 2698
Db 185 CCG-TTCGGTTGGCAGAGCTATGAACAGATATGGGCTGAATACAATCACAGAAATCGTC 243
QY 2699 GTATGCAGTGAACAACTCTCTTTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTATCCG 2758
Db 244 GTATGCAGTGAACAACTCTC-TTCAATTTCTTTATGCGCGGTGTTGGGCGCGTTATTAT-CG 301
```

QY 2759 GAGTTGCAGTTGCGCC 2775
|||||
Db 302 GAGTTGCAGTTGCGCC 318

RESULT 10
CB546382 659 bp mRNA linear EST 01-APR-2003
AMGNNUC:NRHY4-00131-H3-A W Rat hypothalamus (10464) Rattus
norvegicus cDNA clone nrhy4-00131-h3 5', mRNA sequence.

ACCESSION CB546382
VERSION
KEYWORDS
SOURCE
ORGANISM
Rattus norvegicus (Norway rat)

REFERENCE 1 (bases 1 to 659)
AUTHORS Angen EST Program.
TITLE Angen Rat EST Program
JOURNAL Unpublished (2003)
COMMENT Contact: Dan Fitzpatrick
Angen, Inc
One Amgen Center Drive, Thousand Oaks, CA 91320-1799, USA
Tel: 805 447-4881
Plate: 00131 row: h column: 3.

FEATURES
source
1..659
Location/Qualifiers
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone="nrhy4-00131-h3"
/clone_lib="W Rat hypothalamus (10464)"
/note="Vector: pSPORII; Site 1: SalI; Site 2: NotI; W Rat
hypothalamus adult female Wistar rat avg. insert size 2.3
kb fraction 6 and 7"

ORIGIN
Query Match 9.1%; Score 253.8; DB 6; Length 659;
Best Local Similarity 93.7%; Pred. No. 3e-43;
Matches 297; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2459 AAATGAACGCCAAACATAGAAAGGCCCGCGCATTCCTCTAGAGGATG 2518
|||||
Db 5 AAATGAACGCCAAACATAGAAAGGCCCGCGCATTCCTCTAGAGGATG 64

QY 2519 GAACCGCTGGAGAGCAACTGCATAGGCTAGAAAGATACGCCCTGTTCTTGGAAACAA 2578
|||||
Db 65 GAACCGCTGGAGAGCAACTGCATAGGCTATGAAGAGATACGCCCTGTTCTTGGAAACAA 124

QY 2579 TTGCTTTTACAGATGCACATATCGAGTGAAATCACGTTCCGGGAATCTTGAATGT 2638
|||||
Db 125 TTGCTTTTACAGATGCACATATCGAGTGAAATCACGTTCCGGGAATCTTGAATGT 184

QY 2639 CGGTTTCGGTTGGCAGAACCTATGAAACGATATGGCTGAATACAAATCACAGAAATCGTC 2698
|||||
Db 185 CCG-TTCGGTTGGCAGAACCTATGAAACGATATGGCTGAATACAAATCACAGAAATCGTC 243

QY 2699 GTATGCAGTGAACACTCTCTTTCAATTTCTTTATGCGGTTGGGCCCGTTATTTATCCG 2758
|||||
Db 244 GTATGCAGTGAACACTCTC-TTCAATTTCTTTATGCGGTTGGGCCCGTTATTTAT-CG 301

QY 2759 GAGTTGCAGTTGCGCC 2775
|||||
Db 302 GAGTTGCAGTTGCGCC 318

RESULT 11
BZ701200 973 bp DNA linear GSS 19-FEB-2003
LOCUS PUBMP50TD ZM 0.6.1.0_KB Zea mays genomic clone ZMMBta089J04,
DEFINITION genomic survey sequence.

ACCESSION BZ701200
VERSION BZ701200.1 GI:28421047
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Zea.
REFERENCE 1 (bases 1 to 973)
AUTHORS Whitelaw,C.A., Quackenbush,J., Van Aken,S., Utterback,T.,
Resnick,A., Fraser,C.M., Yuan,Y., San Miguel,P., Ma,J. and
Bennetzen,J.
TITLE Maize Genomics Consortium
JOURNAL Unpublished (2003)
COMMENT Contact: Cathy Whitelaw
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA
Tel: 301-838-5843
Fax: 301-838-0208
Email: whitelaw@tigr.org
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..973
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZMMBta089J04"
/clone_lib="ZM 0.6.1.0_KB"
/note="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high
CoT selected genomic DNA library"

ORIGIN
Query Match 9.1%; Score 252.4; DB 8; Length 973;
Best Local Similarity 92.5%; Pred. No. 6.4e-43;
Matches 298; Conservative 0; Mismatches 21; Indels 3; Gaps 3;

QY 2454 TGGTAAATGGAAGACGCCAAACATAGAAAGGCCCGCGCATTCCTCTCTAGA 2513
|||||
Db 8 TGGTAAATGGAAGACGCCAAACATAGAAAGGCCCGCGCATTCCTCTCTAGA 67

QY 2514 GGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAAAGATACGCCCTGTTCTCTGG 2573
|||||
Db 68 GGATGGAACCGCTGGAGAGCAACTGCATAGGCTATGAAGATACGCCCTGTTCTCTGG 127

QY 2574 AACAAATGCTTTTACAGATGCACATATCGAGTGAAATCACGTTCCGGGAATACCTCGA 2633
|||||
Db 128 AACAAATGCTTTTACAGATGCACATATCGAGTGAAATCACGTTCCGGGAATACCTCGA 187

QY 2634 AATGTCCTTTTCGGTTGGCAGAACGCTATGAAACGATATGGCTGAATACAAATCACAGAA 2693
|||||
Db 188 AATGTCCTTTTCGGTTGGCAGAACGCTATGAAACGATATGGCTGAATACAAATCACAGAA 246

QY 2694 TCGTCGTATGCAGTGAAACACTCTCTTTCAATTTCTTTATGCGGCTGTTGGGCCCGTTATTT 2753
|||||
Db 247 TCGTCGTATGCAGTGAAACACTCTC-TTCAATTTCTTTATGCGGCTGTTGGGCCCGTTATTT 305

QY 2754 ATCCGAGTTGCAGTTGCGCC 2775
|||||
Db 306 AT-CCGAGTTGCAGTTGCGCCC 326

RESULT 12
BZ670825 289 bp DNA linear GSS 05-FEB-2003
LOCUS PUBHY64TD ZM 0.6.1.0_KB Zea mays genomic clone ZMMBta059L07,
DEFINITION genomic survey sequence.
ACCESSION BZ670825
VERSION BZ670825.1 GI:28218461
KEYWORDS GSS.
SOURCE Zea mays
ORGANISM Zea mays

Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD Clade; Panicoideae; Andropogoneae; Zea.

REFERENCE

AUTHORS

Whitelaw, C.A., Quackenbush, J., Van Aken, S., Utterback, T., Benneken, J., Fraser, C.M., Yuan, Y., San Miguel, P., Ma, J. and Bannet, J.

TITLE

Maize Genomics Consortium

COMMENT

Unpublished (2003)

Contact: Cathy Whitelaw

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA

Tel: 301-838-5843

Fax: 301-838-0208

Email: whitelaw@tigr.org

Seq primer: TP

Class: sheared ends.

FEATURES

source

1. 289
/organism="Zea mays"
/mol_type="genomic DNA"
/strain="B73"
/db_xref="taxon:4577"
/clone="ZM87A059L07"
/clone_lib="ZM 0.6-1.0 KB"
/notes="Vector: pCR4-TOPO; Site 1: EcoRI; 0.6-1.0 kb high Cot selected genomic DNA library"

ORIGIN

Query Match 8.8% Score 244.6; DB 8; Length 289;
Best Local Similarity 94.5%; Pred. No. 2.5e-41;
Matches 275; Conservative 0; Mismatches 14; Indels 2; Gaps 2;
QY 2445 CGGTACTGTTGGTAAATGAAGACGCGCAAAACATAAAGAGCGCGCCATTCTA 2504
Db 289 CGGTACTGTTGGTAAATGAAGACGCGCAAAACATAAAGAGCGCGCCATTCTA 230
QY 2505 TCCTCTAGAGGATGGAACCGCTGGAGCGCAACTGCATTAAGGCTAGAGATACGCCCT 2564
Db 229 TCCTCTAGAGGATGGAACCGCTGGAGCGCAACTGCATTAAGGCTAGAGATACGCCCT 170
QY 2565 GGTTCCTGGAACATGCTTTTACAGATGCACATATCGAGTGAACATCATCGTTCCGGGA 2624
Db 169 GGTTCCTGGAACATGCTTTTACAGATGCACATATCGAGTGAACATCATCGTTCCGGGA 110
QY 2625 ATACTTCGAATGTCCTGTTGGCAGAGCTATGAACGATATGGCTGAATACAA 2684
Db 109 ATACTTCGAATGTCCTGTTGGCAGAGCTATGAACGATATGGCTGAATACAA 51
QY 2685 ATCAGAGATCGTCGATGTCAGTGAACAACTCTCTTTCAATCTTTATGCGG 2735
Db 50 ATCAGAGATCGTCGATGTCAGTGAACAACTCTCTTTCAATCTTTATGCGG 1

RESULT 13

BY752869

LOCUS

DEFINITION BY752869 RIKEN full-length enriched, adult inner ear Mus musculus

cDNA clone F930028F03 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 697)
Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S., Nikaido, I., Osato, N., Saito, R., Suzuki, H., Yamana, I., Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A., Schonbach, C., Gojobori, T., Baldarelli, R., Hill, D. P., Bult, C., Hume, D. A., Quackenbush, J., Schriml, L. M., Kanapin, A., Matsuda, H., Batalov, S., Beisel, K. W., Blake, J. A., Bradt, D., Brusic, V.,

Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragani, T. A., Fletcher, C. P., Forrest, A., Fraser, K. S., Gaasterland, T., Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grilmond, S., Gustincich, S., Hirokawa, N., Jackson, I. J., Jarvis, E. D., Kanai, A., Kawaji, H., Kawasawa, Y., Kedzierski, R. M., King, B. L., Konagaya, A., Kurochkin, I. V., Lee, Y., Lenhard, B., Lyons, P. A., Maglott, D. R., Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T., Numata, K., Okido, T., Pavan, W. J., Perte, G., Pesole, G., Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S., Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M., Sandelin, A., Schneider, C., Semple, C. A., Setou, M., Shimada, K., Sultana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, M., Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y., Wells, C., Wilming, L. G., Wynshaw-Boris, A., Yanagisawa, M., Yang, I., Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P., Hayatsu, N., Hirozane-Kishikawa, T., Konno, H., Nakamura, M., Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, D., Shibata, K., Arakawa, T., Fukuda, S., Hara, A., Hashizume, W., Imotani, K., Ishii, Y., Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S., Rogers, J., Birney, E. and Hayashizaki, Y.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)

22354683

12466851

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resesgsc.riken.jp, URL: http://genome.gsc.riken.jp/

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,

Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,

Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,

Kondo, S., Konno, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,

Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Saito, R., Sakazume, N.,

Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,

Takeda, Y., Waki, K., Watahiki, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Computational Analysis of Full-Length Mouse cDNAs Compared with

Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multicapillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken

Genomic Sciences Center and Genome Science Laboratory in RIKEN.

Division of Experimental Animal Research in Riken contributed to

prepare mouse tissues.

Tissues were provided by Kirk W. Beisel (Boys Town National

Research Hospital 555 North 30th Street Omaha, NE 68131 USA) whose

assistance we gratefully acknowledge.

Please visit our web site (http://genome.gsc.riken.go.jp) for

further details.

FEATURES

source

1. .697
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="F930028F03"
/tissue_type="inner ear"
/dev_stage="adult"
/clone_lib="RIKEN full-length enriched, adult inner ear"

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2005, 06:54:44 ; Search time 455 Seconds
(without alignments)
10011.856 Million cell updates/sec

Title: US-09-068-751-1
Perfect score: 2784
Sequence: 1 gaagtgaatctgaataatt.....cagttgcgcgcgcgaaca 2784

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/prodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/prodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/prodata/1/ina/PCUS_COMB.seq.*
6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	283.8	10.2	5819	2	US-08-536-559A-2
2	283.8	10.2	5819	3	US-08-862-431-27
3	283.8	10.2	5819	3	US-08-862-431-28
4	283.8	10.2	6092	2	US-08-536-559A-1
5	283.8	10.2	6092	3	US-08-862-431-26
6	280.2	10.1	11616	1	US-08-196-259-2
7	279.8	10.1	5789	3	US-08-862-431-32
8	279.8	10.1	5791	3	US-08-862-431-31
9	279.8	10.1	5793	3	US-08-862-431-29
10	279.8	10.1	5793	3	US-08-862-431-30
11	279.8	10.1	5818	2	US-08-536-559A-3
12	279.8	10.1	5938	2	US-08-536-559A-4
13	279.8	10.1	8937	4	US-08-872-733A-8
14	279.8	10.1	8937	4	US-08-872-733A-9
15	278.6	10.0	5157	2	US-08-474-169-7
16	278.4	10.0	1811	3	US-08-867-352-22
17	278.4	10.0	1838	3	US-08-487-183A-9
18	276.8	9.9	5620	2	US-08-793-170-21
19	276.8	9.9	5620	2	US-08-892-873-21
20	276.8	9.9	5620	3	US-09-334-765A-21
21	276.8	9.9	5620	3	US-09-356-575E-21
22	276.8	9.9	5620	3	US-09-333-820-21
23	276.8	9.9	5620	4	US-09-506-548A-21
24	276.8	9.9	5620	4	US-09-912-552-21
25	276.8	9.9	5620	4	US-08-918-029-21
26	256.4	9.2	12695	4	US-09-949-016-16775
27	254.6	9.1	1798	3	US-09-446-402A-16

28	253.8	9.1	1722	3	US-08-718-425-1	Sequence 1, Appli
29	253.8	9.1	1722	3	US-08-875-277A-1	Sequence 1, Appli
30	253.8	9.1	1722	3	US-09-380-061B-1	Sequence 1, Appli
31	251.8	9.0	1650	1	US-08-354-240A-1	Sequence 1, Appli
32	251.8	9.0	2445	1	US-08-122-520C-8	Sequence 8, Appli
33	251.8	9.0	4544	3	US-09-488-270A-1	Sequence 1, Appli
34	251.8	9.0	5427	3	US-09-282-996-2	Sequence 2, Appli
35	251.2	9.0	6044	1	US-08-316-950-18	Sequence 18, Appli
36	251.2	9.0	6044	5	PCT-US95-12642-18	Sequence 12, Appli
37	251.2	9.0	6047	1	US-08-316-950-12	Sequence 12, Appli
38	251.2	9.0	6047	5	PCT-US95-12642-12	Sequence 12, Appli
39	251.2	9.0	6171	1	US-08-316-950-17	Sequence 17, Appli
40	251.2	9.0	6171	5	PCT-US95-12642-17	Sequence 17, Appli
41	250.2	9.0	10160	3	US-09-097-319A-8	Sequence 8, Appli
42	250.2	9.0	10160	4	US-09-643-971-8	Sequence 8, Appli
43	250.2	9.0	11784	3	US-09-097-319A-9	Sequence 9, Appli
44	250.2	9.0	11784	4	US-09-643-971-9	Sequence 9, Appli
45	250.2	9.0	11991	3	US-09-097-319A-10	Sequence 10, Appli

ALIGNMENTS

RESULT 1
US-08-536-559A-2
; Sequence 2, Application US/08536559A
; Patent No. 5994061
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Constructs and Methods for Screening for
; TITLE OF INVENTION: Increased Expression of Human apo AI Gene
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/536,559A
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5819 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA (plasmid) pGL2(apo
; DESCRIPTION: AI-250) luc"
US-08-536-559A-2

Query Match	10.2%	Score	283.8	DB 2	Length	5819			
Best Local Similarity	90.6%	Pred. No.	3.2e-53						
Matches	336	Conservative	0	Mismatches	32	Indels	3	Gaps	3
Qy	2405	AGCAGGCTCTTGAATTGACCCCTTCGAGCTTGGCATTCGGTACTGTTGGTAAATGG	2464						
Db	242	AGCTGCTCTTAGTCGAGATCTAAGTAAGCTTGGCATTCGGTACTGTTGGTAAATGG	301						
Qy	2465	AGACGCCAAAAACATAAGAAAGGCCCGCGGCATTCTATCTCTAGAGGATGGAAACCG	2524						
Db	302	AGACGCCAAAAACATAAGAAAGGCCCGCGGCATTCTATCTCTAGAGGATGGAAACCG	361						
Qy	2525	CTGGAGAGCCAACTGCATTAAGGCTAGAAGAGATACGCCCTTGTTCTCTGGAAACAATTCGCTT	2584						
Db	362	CTGGAGAGCAACTGCATAAGGCTATGAAGAGATACGCCCTTGTTCTCTGGAAACAATTCGCTT	421						
Qy	2585	TTACAGATGCATATCGAGGTGAACAATCACGTTCCGGGAATATCTTCGAAATTCGCGTTT	2644						
Db	422	TTACAGATGCATATCGAGGTGAACAATCACGTTCCGGGAATATCTTCGAAATTCGCG-TT	480						
Qy	2645	CGGTTGGCAGACCTATGAACGATATGGCTGAATACAAATCAGAGATCGTCGTATGC	2704						
Db	481	CGGTTGGCAGACCTATGAACGATATGGCTGAATACAAATCAGAGATCGTCGTATGC	540						

QY 2705 AGTGAAGAACTCTCTTTCAATTTATGCGCGTGTGGCCCGTTATTTATCCGAGTTG 2764
|||||
Db 541 AGTGAAGAACTCTC-TTCAATTTATGCGCGTGTGGCCCGTTATTTAT-CCGAGTTG 598
QY 2765 CAGTTGCCGCC 2775
|||||
Db 599 CAGTTGCCGCC 609

RESULT 2

US-08-862-431-27
; Sequence 27, Application US/08862431
; Patent No. 6120994
; GENERAL INFORMATION:
; APPLICANT: TAM, SHUI-PANG
; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,431
; FILING DATE: 23-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kim, Judith U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 1669.0020000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5819 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
US-08-862-431-27

Query Match 10.2%; Score 283.8; DB 3; Length 5819;
Best Local Similarity 90.6%; Pred. No. 3.2e-53;
Matches 336; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2405 AGCAGGCTCTTTGAATTCGACCCCTTCGAGCTTGGCAATTCGCGTACTGTGGTAAATGG 2464
Db 242 AGCTGGCTCTTAGTCGAGATCTAAGTAAGCTTGGCAATTCGCGTACTGTGGTAAATGG 301
QY 2465 AAGACGCCAAAACATAAAGAAAGCCCGCGCCCATTTCTATCCTCTAGAGGATGGAACCG 2524
Db 302 AAGACGCCAAAACATAAAGAAAGCCCGCGCCCATTTCTATCCTCTAGAGGATGGAACCG 361
QY 2525 CTGGAGAGCAACTGCATAAGGCTAGAAGAGATAGCCCTGGTTCTCTGGAACAAATTGCTT 2584
Db 362 CTGGAGAGCAACTGCATAAGGCTAGAAGAGATAGCCCTGGTTCTCTGGAACAAATTGCTT 421
QY 2585 TTACAGATGCACATATCGAGGTGAACATCACGTTCCGCGGAATACITTCGAAATGTCGGTTT 2644
Db 422 TTACAGATGCACATATCGAGGTGAACATCACGTTACGCGGAATACITTCGAAATGTCG-IT 480
QY 2645 CGGTTGGCAGAAGCTATGAACGATATGGCTGGAATACAAATACAGAAATCGTCGTATGC 2704
Db 481 CGGTTGGCAGAAGCTATGAACGATATGGCTGGAATACAAATACAGAAATCGTCGTATGC 540

QY 2705 AGTGAAGAACTCTCTTTCAATTTATGCGCGTGTGGCCCGTTATTTATCCGAGTTG 2764
|||||
Db 541 AGTGAAGAACTCTC-TTCAATTTATGCGCGTGTGGCCCGTTATTTAT-CCGAGTTG 598
QY 2765 CAGTTGCCGCC 2775
|||||
Db 599 CAGTTGCCGCC 609

RESULT 3

US-08-862-431-28
; Sequence 28, Application US/08862431
; Patent No. 6120994
; GENERAL INFORMATION:
; APPLICANT: TAM, SHUI-PANG
; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,431
; FILING DATE: 23-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kim, Judith U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 1669.0020000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5819 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
US-08-862-431-28

Query Match 10.2%; Score 283.8; DB 3; Length 5819;
Best Local Similarity 90.6%; Pred. No. 3.2e-53;
Matches 336; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2405 AGCAGGCTCTTTGAATTCGACCCCTTCGAGCTTGGCAATTCGCGTACTGTGGTAAATGG 2464
Db 242 AGCTGGCTCTTAGTCGAGATCTAAGTAAGCTTGGCAATTCGCGTACTGTGGTAAATGG 301
QY 2465 AAGACGCCAAAACATAAAGAAAGCCCGCGCCCATTTCTATCCTCTAGAGGATGGAACCG 2524
Db 302 AAGACGCCAAAACATAAAGAAAGCCCGCGCCCATTTCTATCCTCTAGAGGATGGAACCG 361
QY 2525 CTGGAGAGCAACTGCATAAGGCTAGAAGAGATAGCCCTGGTTCTCTGGAACAAATTGCTT 2584
Db 362 CTGGAGAGCAACTGCATAAGGCTAGAAGAGATAGCCCTGGTTCTCTGGAACAAATTGCTT 421
QY 2585 TTACAGATGCACATATCGAGGTGAACATCACGTTCCGCGGAATACITTCGAAATGTCGGTTT 2644
Db 422 TTACAGATGCACATATCGAGGTGAACATCACGTTACGCGGAATACITTCGAAATGTCG-IT 480
QY 2645 CGGTTGGCAGAAGCTATGAACGATATGGCTGGAATACAAATACAGAAATCGTCGTATGC 2704
Db 481 CGGTTGGCAGAAGCTATGAACGATATGGCTGGAATACAAATACAGAAATCGTCGTATGC 540
QY 2705 AGTGAAGAACTCTCTTTCAATTTATGCGCGTGTGGCCCGTTATTTATCCGAGTTG 2764

Db 541 AGTGAACCTCTC-TTCAATCTTTATGCCGGTGTGGCGCGTTATTTAT-CGGAGTTG 598
Qy 2765 CAGTTGCCGCC 2775
Db 599 CAGTTGCCGCC 609

RESULT 4

US-08-536-559A-1
; Sequence 1, Application US/08536559A
; Patent No. 5994061
; GENERAL INFORMATION:
; APPLICANT: TAM, SHUI-PANG
; TITLE OF INVENTION: DNA Constructs and Methods for Screening for
; TITLE OF INVENTION: Increased Expression of Human apo A1 Gene
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/536,559A
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA (plasmid) pGL2(apo
; DESCRIPTION: AI-491)luc"
US-08-536-559A-1

Query Match 10.2%; Score 283.8; DB 2; Length 6092;
Best Local Similarity 90.6%; Pred. No. 3.2e-53;
Matches 336; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
Qy 2405 AGCAGGCTCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTCTGTTGGTAAATGG 2464
Db 515 AGCTGGCTGCTTAGTCGAGATCTAAGTAAGCTTGGCATTCGGTACTCTGTTGGTAAATGG 574
Qy 2465 AAGACGCCAAAACATAAAGAAAGCCCGCGGCATCTCTATCTCTAGAGGATGGAACCG 2524
Db 575 AAGACGCCAAAACATAAAGAAAGCCCGCGGCATCTCTATCTCTAGAGGATGGAACCG 634
Qy 2525 CTGGAGAGCAACTGCATAGGCTATGAAGAGATACGCCCTGGTTCCTGGAAACAATTGCTT 2584
Db 635 CTGGAGAGCAACTGCATAGGCTATGAAGAGATACGCCCTGGTTCCTGGAAACAATTGCTT 694
Qy 2585 TTACAGATGCACATATCGAGTGAACATCACGTTCCGGGAATACTTTCGAAATGTCCGTTT 2644
Db 695 TTACAGATGCACATATCGAGTGAACATCACGTTCCGGGAATACTTTCGAAATGTCCG-TT 753
Qy 2645 CGTTTGGCAGAGCTATGAACCGATATGGCTGAATACAAATCACAGAATCGTCGTATGC 2704
Db 754 CGTTTGGCAGAGCTATGAACCGATATGGCTGAATACAAATCACAGAATCGTCGTATGC 813
Qy 2705 AGTGAACACTCTTTCAATCTTTATGCCGGTGTGGCGCGCTTATTTATCCGGAGTTG 2764
Db 814 AGTGAACACTCTC-TTCAATCTTTATGCCGGTGTGGCGCGCTTATTTAT-CGGAGTTG 871
Qy 2765 CAGTTGCCGCC 2775
Db 872: CAGTTGCCGCC 882

RESULT 5

US-08-862-431-26
; Sequence 26, Application US/08862431
; Patent No. 6120994
; GENERAL INFORMATION:

; APPLICANT: TAM, SHUI-PANG
; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,431
; FILING DATE: 23-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kim, Judith U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 1669.0020000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 26:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6092 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
US-08-862-431-26

Query Match 10.2%; Score 283.8; DB 3; Length 6092;
Best Local Similarity 90.6%; Pred. No. 3.2e-53;
Matches 336; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
Qy 2405 AGCAGGCTCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTCTGTTGGTAAATGG 2464
Db 515 AGCTGGCTGCTTAGTCGAGATCTAAGTAAGCTTGGCATTCGGTACTCTGTTGGTAAATGG 574
Qy 2465 AAGACGCCAAAACATAAAGAAAGCCCGCGGCATCTCTATCTCTAGAGGATGGAACCG 2524
Db 575 AAGACGCCAAAACATAAAGAAAGCCCGCGGCATCTCTATCTCTAGAGGATGGAACCG 634
Qy 2525 CTGGAGAGCAACTGCATAGGCTATGAAGAGATACGCCCTGGTTCCTGGAAACAATTGCTT 2584
Db 635 CTGGAGAGCAACTGCATAGGCTATGAAGAGATACGCCCTGGTTCCTGGAAACAATTGCTT 694
Qy 2585 TTACAGATGCACATATCGAGTGAACATCACGTTCCGGGAATACTTTCGAAATGTCCGTTT 2644
Db 695 TTACAGATGCACATATCGAGTGAACATCACGTTCCGGGAATACTTTCGAAATGTCCG-TT 753
Qy 2645 CGTTTGGCAGAGCTATGAACCGATATGGCTGAATACAAATCACAGAATCGTCGTATGC 2704
Db 754 CGTTTGGCAGAGCTATGAACCGATATGGCTGAATACAAATCACAGAATCGTCGTATGC 813
Qy 2705 AGTGAACACTCTTTCAATCTTTATGCCGGTGTGGCGCGCTTATTTATCCGGAGTTG 2764
Db 814 AGTGAACACTCTC-TTCAATCTTTATGCCGGTGTGGCGCGCTTATTTAT-CGGAGTTG 871
Qy 2765 CAGTTGCCGCC 2775
Db 872: CAGTTGCCGCC 882

RESULT 6

US-08-196-259-2
; Sequence 2, Application US/08196259
; Patent No. 5639596
; GENERAL INFORMATION:
; APPLICANT: BORNKAMM, Georg

APPLICANT: POLACK, Axel
TITLE OF INVENTION: DNA CONSTRUCT AND IN VITRO TEST FOR
TITLE OF INVENTION: DETECTING TUMOR PROMOTERS BY MEANS OF SAID DNA CONSTRUCT
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: SPENCER, FRANK & SCHNEIDER
STREET: Suite 300 East - 1100 New York Ave., N.W.
CITY: Washington
STATE: DC
COUNTRY: US
ZIP: 20005-3955
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/196,259
FILING DATE: 25-JUL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Spencer, George H.
REGISTRATION NUMBER: 18,038
REFERENCE/DOCKET NUMBER: KFK 0449
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 414-4000
TELEFAX: (202) 414-4040
TELEX: 64267
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 11616 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PUBLICATION INFORMATION:
DOCUMENT NUMBER: WO PCT/DE91/00652
FILING DATE: 16-AUG-1991
PUBLICATION DATE: 04-MAR-1993
US-08-196-259-2

Query Match 10.1%; Score 280.2; DB 1; Length 11616;
Best Local Similarity 93.9%; Pred. No. 2.5e-52;
Matches 324; Conservative 0; Mismatches 18; Indels 3; Gaps 3;
QY 2431 CGAGCTTGGCAATTCGGGTACTGTGGTAAATGGAGAGCGCCAAACATAAAGAAAGGC 2490
DB 1963 CAAGCTTGGCAATTCGGGTACTGTGGTAAATGGAGAGCGCCAAACATAAAGAAAGGC 2022
QY 2491 CGGGGCCAATTCCTCTAGAGGATGGAACCGCTGGAGAGCCAACTGCATAAGGCTAG 2550
DB 2023 CCGGGGCCAATTCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATG 2082
QY 2551 AAGAGATACGCTGTCTCGAACAATGCTTTTACAGATGCACATATCGAGGTGAAC 2610
DB 2083 AAGAGATACGCTGTCTCGAACAATGCTTTTACAGATGCACATATCGAGGTGAAC 2142
QY 2611 ATCAGCTTCGCGGAATFACATTCGAAATGTCGCTTTTGGTGGCAGAGCTATGAAACGATA 2670
DB 2143 ATCAGCTTCGCGGAATFACATTCGAAATGTCGCTTTTGGTGGCAGAGCTATGAAACGATA 2201
QY 2671 TGGGCTGAATACAAATACAGAAATCGTCTGATGCAAGTGAACCTCTCTTCAATCTTTTA 2730
DB 2202 TGGGCTGAATACAAATACAGAAATCGTCTGATGCAAGTGAACCTCTCTTCAATCTTTTA 2260
QY 2731 TCGCGGTGTGGCGCGGTATTATATCCGGAGTTGCAGTTGCGGCC 2775
DB 2261 TCGCGGTGTGGCGCGGTATTATATCCGGAGTTGCAGTTGCGGCC 2304

RESULT 7

US-08-862-431-32
Sequence 32, Application US/08862431
Patent No. 6120994
GENERAL INFORMATION:
APPLICANT: TAM, SHUI-PANG
TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
NUMBER OF SEQUENCES: 51
CORRESPONDENCE ADDRESS:
ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
STREET: 1100 NEW YORK AVENUE, SUITE 600
CITY: WASHINGTON
STATE: DC
COUNTRY: US
ZIP: 20005-3934
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/862,431
FILING DATE: 23-MAY-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Kim, Judith U.
REGISTRATION NUMBER: 40,679
REFERENCE/DOCKET NUMBER: 1669.0020000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 371-2600
TELEFAX: (202) 371-2540
INFORMATION FOR SEQ ID NO: 32:
SEQUENCE CHARACTERISTICS:
LENGTH: 5789 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
US-08-862-431-32

Query Match 10.1%; Score 279.8; DB 3; Length 5789;
Best Local Similarity 94.2%; Pred. No. 2.5e-52;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 2433 AGCTTGGCAATTCGGGTACTGTGGTAAATGGAGAGCGCCAAACATAAAGAAAGGCC 2492
DB 240 AGCTTGGCAATTCGGGTACTGTGGTAAATGGAGAGCGCCAAACATAAAGAAAGGCC 299
QY 2493 GCGGCCAATTCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTAGAA 2552
DB 300 GCGGCCAATTCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAA 359
QY 2553 GAGATACGCTGTCTCGAACAATGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
DB 360 GAGATACGCTGTCTCGAACAATGCTTTTACAGATGCACATATCGAGGTGAACAT 419
QY 2613 CACGTTTCGCGGAATFACATTCGAAATGTCGCTTTTGGTGGCAGAGCTATGAAACGATG 2672
DB 420 CACGTTTCGCGGAATFACATTCGAAATGTCGCTTTTGGTGGCAGAGCTATGAAACGATG 478
QY 2673 GGCTGAATACAAATACAGAAATCGTCTGATGCAAGTGAACCTCTCTTCAATCTTTTATG 2732
DB 479 GGCTGAATACAAATACAGAAATCGTCTGATGCAAGTGAACCTCTCTTCAATCTTTTATG 537
QY 2733 CCGGTGTGGCGCGGTATTATATCCGGAGTTGCAGTTGCGGCC 2775
DB 538 CCGGTGTGGCGCGGTATTATATCCGGAGTTGCAGTTGCGGCC 579

RESULT 8
US-08-862-431-31
Sequence 31, Application US/08862431
Patent No. 6120994
GENERAL INFORMATION:
APPLICANT: TAM, SHUI-PANG

;; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
;; NUMBER OF SEQUENCES: 51
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
;; STREET: 1100 NEW YORK AVENUE, SUITE 600
;; CITY: WASHINGTON
;; STATE: DC
;; COUNTRY: US
;; ZIP: 20005-3934
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 23-MAY-1997
;; APPLICATION NUMBER: US/08/862,431
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kim, Judith U.
;; REGISTRATION NUMBER: 40,679
;; REFERENCE/DOCKET NUMBER: 1669.0020000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 371-2600
;; TELEFAX: (202) 371-2540
;; INFORMATION FOR SEQ ID NO: 31:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5791 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: circular
;; US-08-862-431-31

Query Match 10.1%; Score 279.8; DB 3; Length 5791;
Best Local Similarity 94.2%; Pred. No. 2.5e-52;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 2433 AGCTTGGCAATTCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATAAAGAAAGGCC 2492
DB 242 AGCTTGGCAATTCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATAAAGAAAGGCC 301
QY 2493 GGGCCATCTTATCTCTAGAGATGGACCGCTGGAGAGCAACTGCATGAAGCTAGAA 2552
DB 302 GGGCCATCTTATCTCTAGAGATGGACCGCTGGAGAGCAACTGCATGAAGCTAGAA 361
QY 2553 GAGATACGCCCTGTTCTCGAACAATGCTTTTACAGATGCACATATCGAGTGAACAT 2612
DB 362 GAGATACGCCCTGTTCTCGAACAATGCTTTTACAGATGCACATATCGAGTGAACAT 421
QY 2613 CACGTTGCGGGAATCTTCGAAATGTCGGTTTCGGTTGGCAGAACTATGAAACGATATG 2672
DB 422 CACGTTGCGGGAATCTTCGAAATGTCGGTTTCGGTTGGCAGAACTATGAAACGATATG 480
QY 2673 GCGTGAATACAAATCAGAAATCGTATGCAAGTGAAGTCTCTTCAATTTCTTTATG 2732
DB 481 GCGTGAATACAAATCAGAAATCGTATGCAAGTGAAGTCTCTTCAATTTCTTTATG 539
QY 2733 CCGGTGTGGCGCGGTTATTATTCGGAGTTGCGAGTTGCGGCC 2775
DB 540 CCGGTGTGGCGCGGTTATTATTCGGAGTTGCGAGTTGCGGCC 581

RESULT 9
US-08-862-431-29
; Sequence 29, Application US/08862431
; Patent No. 6120994
; GENERAL INFORMATION:
; APPLICANT: TAM, SHUI-PANG
; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600

;; CITY: WASHINGTON
;; STATE: DC
;; COUNTRY: US
;; ZIP: 20005-3934
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent in Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; FILING DATE: 23-MAY-1997
;; APPLICATION NUMBER: US/08/862,431
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Kim, Judith U.
;; REGISTRATION NUMBER: 40,679
;; REFERENCE/DOCKET NUMBER: 1669.0020000
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (202) 371-2600
;; TELEFAX: (202) 371-2540
;; INFORMATION FOR SEQ ID NO: 29:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 5793 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: double
;; TOPOLOGY: circular
;; US-08-862-431-29

Query Match 10.1%; Score 279.8; DB 3; Length 5793;
Best Local Similarity 94.2%; Pred. No. 2.5e-52;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 2433 AGCTTGGCAATTCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATAAAGAAAGGCC 2492
DB 244 AGCTTGGCAATTCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATAAAGAAAGGCC 303
QY 2493 GGGCCATCTTATCTCTAGAGATGGACCGCTGGAGAGCAACTGCATGAAGCTAGAA 2552
DB 304 GGGCCATCTTATCTCTAGAGATGGACCGCTGGAGAGCAACTGCATGAAGCTAGAA 363
QY 2553 GAGATACGCCCTGTTCTCGAACAATGCTTTTACAGATGCACATATCGAGTGAACAT 2612
DB 364 GAGATACGCCCTGTTCTCGAACAATGCTTTTACAGATGCACATATCGAGTGAACAT 423
QY 2613 CACGTTGCGGGAATCTTCGAAATGTCGGTTTCGGTTGGCAGAACTATGAAACGATATG 2672
DB 424 CACGTTGCGGGAATCTTCGAAATGTCGGTTTCGGTTGGCAGAACTATGAAACGATATG 482
QY 2673 GCGTGAATACAAATCAGAAATCGTATGCAAGTGAAGTCTCTTCAATTTCTTTATG 2732
DB 483 GCGTGAATACAAATCAGAAATCGTATGCAAGTGAAGTCTCTTCAATTTCTTTATG 541
QY 2733 CCGGTGTGGCGCGGTTATTATTCGGAGTTGCGAGTTGCGGCC 2775
DB 542 CCGGTGTGGCGCGGTTATTATTCGGAGTTGCGAGTTGCGGCC 583

RESULT 10
US-08-862-431-30
; Sequence 30, Application US/08862431
; Patent No. 6120994
; GENERAL INFORMATION:
; APPLICANT: TAM, SHUI-PANG
; TITLE OF INVENTION: ANTIOXIDANT RESPONSIVE ELEMENT
; NUMBER OF SEQUENCES: 51
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERNE, KESSLER, GOLDSTEIN & FOX P.L.L.C.
; STREET: 1100 NEW YORK AVENUE, SUITE 600
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20005-3934
; COMPUTER READABLE FORM:

```

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/862,431
; FILING DATE: 23-MAY-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Kim, Judith U.
; REGISTRATION NUMBER: 40,679
; REFERENCE/DOCKET NUMBER: 1669.0020000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 30:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5793 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; US-08-862-431-30

Query Match      10.1%; Score 279.8; DB 3; Length 5793;
Best Local Similarity 94.2%; Pred. No. 2.5e-52;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2433 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGCGCCAAACATATAAGAAAGGCC 2492
Db 244 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGCGCCAAACATATAAGAAAGGCC 303

QY 2493 GCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 2552
Db 304 GCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 363

QY 2553 GAGATACGCCCTGGTTCCTGGAAACAATTCCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 364 GAGATACGCCCTGGTTCCTGGAAACAATTCCTTTTACAGATGCACATATCGAGGTGAACAT 423

QY 2613 CACGTTCCGGGAATCTCCGAAATGTCGTTTCGGTTGGCAGAACGCTATGAACCATATG 2672
Db 424 CACGTTCCGGGAATCTCCGAAATGTCGTTTCGGTTGGCAGAACGCTATGAACCATATG 482

QY 2673 GGCTGAATACAAATCACAGAAATCGTATGCAAGTAAACCTCTCTTTCAATCTTTATG 2732
Db 483 GGCTGAATACAAATCACAGAAATCGTATGCAAGTAAACCTCTCTTTCAATCTTTATG 541

QY 2733 CCGGTGTTGGGCCGTTATTTATCCGAGTTCGAGTTGCGGCC 2775
Db 542 CCGGTGTTGGGCCGTTATTTATCCGAGTTCGAGTTGCGGCC 583

RESULT 11
US-08-536-559A-3
; Sequence 3, Application US/08536559A
; Patent No. 5994061
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Constructs and Methods for Screening for
; TITLE OF INVENTION: Increased Expression of Human apo AI Gene
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/536,559A
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5818 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double

```

```

; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA (plasmid) pGL2(apo
; DESCRIPTION: AI-DRE)TK/luc"
; US-08-536-559A-3

Query Match      10.1%; Score 279.8; DB 2; Length 5818;
Best Local Similarity 94.2%; Pred. No. 2.5e-52;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2433 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGCGCCAAACATATAAGAAAGGCC 2492
Db 269 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGCGCCAAACATATAAGAAAGGCC 328

QY 2493 GCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 2552
Db 329 GCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 388

QY 2553 GAGATACGCCCTGGTTCCTGGAAACAATTCCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 389 GAGATACGCCCTGGTTCCTGGAAACAATTCCTTTTACAGATGCACATATCGAGGTGAACAT 448

QY 2613 CACGTTCCGGGAATCTCCGAAATGTCGTTTCGGTTGGCAGAACGCTATGAACCATATG 2672
Db 449 CACGTTCCGGGAATCTCCGAAATGTCGTTTCGGTTGGCAGAACGCTATGAACCATATG 507

QY 2673 GGCTGAATACAAATCACAGAAATCGTATGCAAGTAAACCTCTCTTTCAATCTTTATG 2732
Db 508 GGCTGAATACAAATCACAGAAATCGTATGCAAGTAAACCTCTCTTTCAATCTTTATG 566

QY 2733 CCGGTGTTGGGCCGTTATTTATCCGAGTTCGAGTTGCGGCC 2775
Db 567 CCGGTGTTGGGCCGTTATTTATCCGAGTTCGAGTTGCGGCC 608

RESULT 12
US-08-536-559A-4
; Sequence 4, Application US/08536559A
; Patent No. 5994061
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DNA Constructs and Methods for Screening for
; TITLE OF INVENTION: Increased Expression of Human apo AI Gene
; NUMBER OF SEQUENCES: 17
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/536,559A
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5938 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: other nucleic acid
; DESCRIPTION: /desc = "DNA (plasmid) pGL2(4x apo
; DESCRIPTION: AI-DRE)TK/luc"
; US-08-536-559A-4

Query Match      10.1%; Score 279.8; DB 2; Length 5938;
Best Local Similarity 94.2%; Pred. No. 2.5e-52;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2433 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGCGCCAAACATATAAGAAAGGCC 2492
Db 389 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGCGCCAAACATATAAGAAAGGCC 448

QY 2493 GCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 2552
Db 449 GCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 508

```


Qy	2553	GAGATACGCCCTGGTTCTCTGGAA	CAATTGCTTTTACAGATG	CACATATCGAGGTGAACAT	26112
Db	509	GAGATACGCCCTGGTTCTCTGGAA	CAATTGCTTTTACAGATG	CACATATCGAGGTGAACAT	568
Qy	2613	CACGTTTCGCGGAATACATTCGAA	ATGTCGGTTTCGGTTGGCAGAGCT	ATGAAACGATATG	26772
Db	569	CACGTAACGCGGAATACATTCGAA	ATGTCGGTTTCGGTTGGCAGAGCT	ATGAAACGATATG	627
Qy	2673	GGCTGAATACAAATACACAGAA	TCTCGTATGCACTGAAATCTCTCT	TTCATTTCTTTATG	2732
Db	628	GGCTGAATACAAATACACAGAA	TCTCGTATGCACTGAAATCTCTC	TTCATTTCTTTATG	686
Qy	2733	CCGGTGTGTGGCCCGTTATTAT	TATCCGAGTTGCAAGTTGCCGCC		2775
Db	687	CCGGTGTGTGGCCCGTTATTAT	TATCCGAGTTGCAAGTTGCCGCC		728

RESULT 13
US-09-872-733A-8
; Sequence 8, Application US/09872733A
; Patent No. 6656706
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, HIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287U51 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872, 733A
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173, 036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBCwCnLuci
US-09-872-733A-8

RESULT 14									
US-09-872-733A-9									
; Sequence 9, Application US/09872733A									
; Patent No. 6656706									
; GENERAL INFORMATION:									
; APPLICANT: The Government of the United States of America, as									
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, HIV GAG AND									
; TITLE OF INVENTION: HIV ENV GENES									
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, HIV GAG & ENV									
; CURRENT APPLICATION NUMBER: US/09/872,733A									
; CURRENT FILING DATE: 2001-06-01									
; PRIOR APPLICATION NUMBER: PCT/US00/34985									
; PRIOR FILING DATE: 2000-12-22									
; PRIOR APPLICATION NUMBER: 60/173,036									
; PRIOR FILING DATE: 1999-12-23									
; NUMBER OF SEQ ID NOS: 19									
; SOFTWARE: PatentIn Ver. 2.1									
; SEQ ID NO 9									
; LENGTH: 8937									
; TYPE: DNA									
; ORGANISM: Artificial Sequence									
; FEATURE:									
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence									
; OTHER INFORMATION: of transfer construct pmBcmCnluci									
US-09-872-733A-9									
Query Match	10.1%;	Score 279.8;	DB 4;	Length 8937;					
Best Local Similarity	94.2%;	Pred. No. 2.8e-52;							
Matches 323;	Conservative 0;	Mismatches 17;	Indels 3;	Gaps 3;					
Qy	2433	AGCTTGCCATTCTCGGTACTCTGTTGTAATGGAAGACGCCAAAACATAAAGAAAGGCC	2492						
Db	1867	AGCTTGCCATTCTCGGTACTCTGTTGTAATGGAAGACGCCAAAACATAAAGAAAGGCC	1926						
Qy	2493	GGCGCCATTCTATCTCTAGAGGATGGAACCGCTGGAGAGCCAACTGCATTAAGGCTAGAA	2552						
Db	1927	GGCGCCATTCTATCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATTAAGGCTAGAA	1986						
Qy	2553	GAGATAGCCCTCGTTCTCTGGAAACAATGCTTTTACAGATGCACATATCAGAGTGAAACAT	2612						
Db	1987	GAGATAGCCCTCGTTCTCTGGAAACAATGCTTTTACAGATGCACATATCAGAGTGAAACAT	2046						
Qy	2613	CAGCTTCGCGGAATACCTTCGAAATGTCGGTTTCGGTTGGCAGAGGCTATCAAAACGATATG	2672						
Db	2047	CAGCTTCGCGGAATACCTTCGAAATGTCGGTTTCGGTTGGCAGAGGCTATCAAAACGATATG	2105						
Qy	2673	GGCTGGAATACAAATACAGAAATCGTCGTATGTCAGTGAAAACTCTCTTTCAATTTCTTTATG	2732						
Db	2106	GGCTGGAATACAAATACAGAAATCGTCGTATGTCAGTGAAAACTCTCTTTCAATTTCTTTATG	2164						
Qy	2733	CCGGTGTTGGCCGGTTATTTATCCGAGGTGAGTTGGCGCC	2775						
Db	2165	CCGGTGTTGGCCGGTTATTTAT - CGAGTTCAGTTGGCGCC	2206						

RESULT 15
US-08-474-169-7
; Sequence 7, Application US/08474169
; Patent NO. 5851796
; GENERAL INFORMATION:
; APPLICANT: Schatz, David G.
; TITLE OF INVENTION: An Autoregulatory Tetracycline-Regulated
; TITLE OF INVENTION: System for Inducible Gene Expression in Eucaryotes
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: STERN, KESSLER, GOLDSTEIN & FOX P.L.L.C
; STREET: 1100 New York Ave. Suite 600
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

Search completed: August 1, 2005, 11:40:23
Job time : 458 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2005, 08:41:06 ; Search time 11842 Seconds
(without alignments)
11391.599 Million cell updates/sec

Title: US-09-068-751-1
Perfect score: 2784
Sequence: 1 gaagtgaatctgaataatt.....cagtgccgcgcgcgaaca 2784

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match	Description
1	2760	99.1	A93968 Sequence 1
2	2054	73.8	AC105794 Rattus no
3	1975.2	70.9	U26708 Rattus norv
4	1959.2	70.4	S80994 MLC-2-myosin
C 5	1473.6	52.9	Continuation (3 of
C 6	1473.6	52.9	Continuation (4 of
7	1338	48.1	AC137306 Rattus no
C 8	1156.2	41.5	AC137290 Rattus no
9	497	17.9	M30298 Rat cardiac
10	437.4	15.7	AF302688 Mus muscu
11	436.4	15.7	AF326789 Mus muscu
12	429.6	15.4	AC140353 Mus muscu
C 13	290.6	10.4	AX11089 Mus muscu
14	284.6	10.2	AX701675 Sequence
15	284.6	10.2	AJ277960 Cloning v
16	284.6	10.2	AJ277959 Cloning v
17	283.8	10.2	AR089206 Sequence
18	283.8	10.2	AR089205 Sequence
19	282.4	10.1	U25266 Ligation-in

20	282.2	10.1	5675	12	AF053698	Reporter
21	282.2	10.1	5688	12	AF053461	Reporter
22	282.2	10.1	5695	12	AF053315	Reporter
23	282.2	10.1	5714	12	AF167311	Reporter
24	282.2	10.1	5743	12	AF053462	Reporter
25	282.2	10.1	5748	12	AF058756	Cloning v
26	282.2	10.1	5909	12	AF081837	Cloning v
27	281.4	10.1	5406	12	AF093688	MMTV-luc1
28	281.4	10.1	5838	12	AF093687	MMTV-luc1
29	281.4	10.1	6505	12	AF093686	MMTV-luc1
30	280.8	10.1	5067	12	AF093685	Luciferase
31	280.8	10.1	5499	12	AF093684	Luciferase
32	280.8	10.1	6163	12	AF093682	Luciferase
33	280.8	10.1	6163	12	AF093682	Luciferase
34	280.2	10.1	5847	12	AF311601	Reporter
35	280.2	10.1	11616	6	A27170	Artificial
36	280.2	10.1	11616	6	A39929	Sequence 1
37	280.2	10.1	11616	6	I45936	Sequence 2
38	279.8	10.1	5598	12	CVFGL2BAS	Cloning vec
39	279.8	10.1	5790	12	CVFGL2PRO	Cloning vec
40	279.8	10.1	5818	6	AR089207	Sequence
41	279.8	10.1	5849	6	Q815105	Sequence
42	279.8	10.1	5855	12	CVFGL2ENH	Cloning vec
43	279.8	10.1	5938	6	AR089208	Sequence
44	279.8	10.1	6047	12	CVFGL2CON	Cloning vec
45	279.8	10.1	8937	6	AR435499	Sequence

ALIGNMENTS

RESULT 1
A93968
LOCUS A93968 2784 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9717937.
ACCESSION A93968
VERSION A93968.1 GI:6742069
KEYWORDS .
ORGANISM unidentified
SOURCE unidentified
REFERENCE 1 (bases 1 to 2784)
AUTHORS Franz, W. and Rothmann, T.
TITLE GENE-THERAPEUTIC NUCLEIC ACID CONSTRUCT, PRODUCTION OF SAME AND USE OF SAME IN THE TREATMENT OF HEART DISORDERS
JOURNAL Patent: WO 9717937-A 1 22-MAY-1997;
FRANZ WOLFGANG M (DE); ROTHMANN THOMAS (DE)
FEATURES
source
1..2784
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 99.1%; Score 2760; DB 6; Length 2784;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 15; Indels 0; Gaps 0;
Qy 1 GAAGTGAATCTCAATAATTTTGTGTACTCATAGCGCGTAATATTGTCTAGGCGCGC 60
Db 1 GAAGTGAATCTCAATAATTTTGTGTACTCATAGCGCGTAATATTGTCTAGGCGCGC 60
Qy 61 GGACTTTGACCGTTACGTGGAGACTCGCCAGGTGTTTTCTCAGGTGTTTTCCCGGTT 120
Db 61 GGACTTTGACCGTTACGTGGAGACTCGCCAGGTGTTTTCTCAGGTGTTTTCCCGGTT 120
Qy 121 CCGGTCAAAGTTGGCGTTTATATATAGTCAGGGGATCCGGAATTTCTTGAAGACGA 180
Db 121 CCGGTCAAAGTTGGCGTTTATATATAGTCAGGGGATCCGGAATTTCTTGAAGACGA 180
Qy 181 AAGGGCCCGGTACCCAGGACTGATCTCGAAAGTTCTAGGCTGCAGAAATCTCACACGC 240
Db 181 AAGGGCCCGGTACCCAGGACTGATCTCGAAAGTTCTAGGCTGCAGAAATCTCACACGC 240

QY	241	ACAAGAGTTTGGAGTCACAGGATGGGTGCTCGC	AAAGAGCCTAGGACAGAAAGCTTTGTCA	300
Db	241	ACNAGAGTTTGGAGTCACAGGATGGGTGCTCGC	CAAGAGCCTAGGACAGAAAGCTTTGTCA	300
QY	301	GCCCTGTGCCCGGACCCCTGTGGACTGTGAGA	AGAGCAGAGTCCCCACCCAGGCCCTTCT	360
Db	301	GCCCTGTGCCCGGACCCCTGTGGACTGTGAGA	AGAGCAGAGTCCCCACCCAGGCCCTTCT	360
QY	361	TAGACCCACCCCGGGTTTTCCAGCATCTTCTG	CAGGACCGGACCCCTGGCTGAAAGT	420
Db	361	TAGACCCACCCCGGGTTTTCCAGCATCTTCTG	CAGGACCGGACCCCTGGCTGAAAGT	420
QY	421	ACAGAAACCTTAGAGTCTGAGGCCATGTGGCT	TGGGCCGCCATGTTTCCAGAAATCCTCTG	480
Db	421	ACAGAAACCTTAGAGTCTGAGGCCATGTGGCT	TGGGCCGCCATGTTTCCAGAAATCCTCTG	480
QY	481	GTCTAAGGATCCAGACCTCTTACGGAGCCCA	CAAGCTCAAGGACAGTTAGCATGTTCA	540
Db	481	GTCTAAGGATCCAGACCTCTTACGGAGCCCA	CAAGCTCAAGGACAGTTAGCATGTTCA	540
QY	541	GTGTACTGGGAGGAGCAGGAGCCCAAGAGTCA	TGAAGTCCACAGGGGCTCCGGTTCC	600
Db	541	GTGTACTGGGAGGAGCAGGAGCCCAAGAGTCA	TGAAGTCCACAGGGGCTCCGGTTCC	600
QY	601	GAGGCCCTTGGGTTTTATCACAAATGTTTTCC	CACCCAGCAACATAAAACAGCTCCTCAG	660
Db	601	GAGGCCCTTGGGTTTTATCACAAATGTTTTCC	CACCCAGCAACATAAAACAGCTCCTCAG	660
QY	661	ACAGCCAGTGGACCAAGTGGACCAAGGACAG	ATCACCTCTGTGGGCCAGACATCATAG	720
Db	661	ACAGCCAGTGGACCAAGTGGACCAAGGACAG	ATCACCTCTGTGGGCCAGACATCATAG	720
QY	721	TAACTCTAACCCTCAATCTCCAGCCTCCAC	AGTCAATGTTCGGTCACTTGTCTCAGC	780
Db	721	TAACTCTAACCCTCAATCTCCAGCCTCCAC	AGTCAATGTTCGGTCACTTGTCTCAGC	780
QY	781	CACCACACTTGGCAAGTCAAGTGTGCTCAAC	ACAATCTTCAGAAGCCAGGGGATGGGG	840
Db	781	CACCACACTTGGCAAGTCAAGTGTGCTCAAC	ACAATCTTCAGAAGCCAGGGGATGGGG	840
QY	841	TTTTGTGTTAACTGATGGGTGTTTTGTTTTG	TTTTGTTTTTTCATTAAGTGTCAAGTGT	900
Db	841	TTTTGTGTTAACTGATGGGTGTTTTGTTTTG	TTTTGTTTTTTCATTAAGTGTCAAGTGT	900
QY	901	GCTAGCTTGAACCTACTATGAGGCAAGCAT	GACCATGAACTTCTGATCCTCCTTCCTC	960
Db	901	GCTAGCTTGAACCTACTATGAGGCAAGCAT	GACCATGAACTTCTGATCCTCCTTCCTC	960
QY	961	AGTGTCTGGGATAACAGGTTGTGTGCTCACT	CCCTACCTTCTAATAGCAATATGTGGCA	1020
Db	961	AGTGTCTGGGATAACAGGTTGTGTGCTCACT	CCCTACCTTCTAATAGCAATATGTGGCA	1020
QY	1021	CATGTTTGTGCCCCACAGGTTGAGACCATCT	TGACCTGAGGAAGAAATAGCTAAACATCA	1080
Db	1021	CATGTTTGTGCCCCACAGGTTGAGACCATCT	TGACCTGAGGAAGAAATAGCTAAACATCA	1080
QY	1081	CCTCTGAAAGTTGCTTGGATCTCGTCTTTGT	CTTTTCCAGCACTCAGGAGTGGGGGGTC	1140
Db	1081	CCTCTGAAAGTTGCTTGGATCTCGTCTTTGT	CTTTTCCAGCACTCAGGAGTGGGGGGTC	1140
QY	1141	AGAAGTGCAAAGTCAGCCCTGCTACATAAT	GAGTTCAAGGCTTCGGCTTACATGAG	1200
Db	1141	AGAAGTGCAAAGTCAGCCCTGCTACATAAT	GAGTTCAAGGCTTCGGCTTACATGAG	1200
QY	1201	ACCATGCTCAAAAGAAAGAAATGGTATAGTG	ACATCTCTGGTCTCTCCAGTACTT	1260
Db	1201	ACCATGCTCAAAAGAAAGAAATGGTATAGTG	ACATCTCTGGTCTCTCCAGTACTT	1260
QY	1261	AGGGACACAGAGGCGACTCCACCACTCTCC	AGCAGCTGGGCTGCCCTCCCGAGCCTCG	1320
Db	1261	AGGGACACAGAGGCGACTCCACCACTCTCC	AGCAGCTGGGCTGCCCTCCCGAGCCTCG	1320

QY	1321	TTTTATTTCATATCAATGAGATGGGACCCAA	CTGCTAAAGTGACCTTGACCCACGGGT	1380
Db	1321	TTTTATTTCATATCAATGAGATGGGACCCAA	CTGCTAAAGTGACCTTGACCCACGGGT	1380
QY	1381	GACTGGAGACTTGAGAGTGGAGGGTTTATCA	TTTTCTCAGTTCGGTCAGCAATGTGTGCGC	1440
Db	1381	GACTGGAGACTTGAGAGTGGAGGGTTTATCA	TTTTCTCAGTTCGGTCAGCAATGTGTGCGC	1440
QY	1441	GCCAGAGAGTTTTCAGTTCAAAGTAGAAGAT	TGGGACAGGGAGACCCAGCAGAGACC	1500
Db	1441	GCCAGAGAGTTTTCAGTTCAAAGTAGAAGAT	TGGGACAGGGAGACCCAGCAGAGACC	1500
QY	1501	CCACCTCGAGCTGACTGTCCCTGTGCGGCT	TGGGTGGGACACAAAGCAGAGAACAGAG	1560
Db	1501	CCACCTCGAGCTGACTGTCCCTGTGCGGCT	TGGGTGGGACACAAAGCAGAGAACAGAG	1560
QY	1561	GCAGAGAACAGGGTGGGTGACATTTGAGCA	AGGATGGGGGTGTCAGAGGCTGCCCAA	1620
Db	1561	GCAGAGAACAGGGTGGGTGACATTTGAGCA	AGGATGGGGGTGTCAGAGGCTGCCCAA	1620
QY	1621	GATGCATAGGTGCAAGGCCCTGAGTTTCG	AGGATGCCCTGGATCCGGAATCAAGAGT	1680
Db	1621	GATGCATAGGTGCAAGGCCCTGAGTTTCG	AGGATGCCCTGGATCCGGAATCAAGAGT	1680
QY	1681	GCTCTCCTCTTCTCCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1740
Db	1681	GCTCTCCTCTTCTCCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1740
QY	1741	TCGCCCCCTCTTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1800
Db	1741	TCGCCCCCTCTTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1800
QY	1801	CCTCTCTCCCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860
Db	1801	CCTCTCTCCCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1860
QY	1861	CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1920
Db	1861	CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1920
QY	1921	TGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1980
Db	1921	TGCCCCCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	1980
QY	1981	TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	2040
Db	1981	TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	2040
QY	2041	CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	2100
Db	2041	CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	2100
QY	2101	CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	2160
Db	2101	CCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	CTCTCTCTCTCTCTCTCTCTCTCTCTCT	2160
QY	2161	GAGCAGAGAGCATCGTTCAGGCCAGGCC	CCAGCCACCTGTCTTAACTTGAAGGCA	2220
Db	2161	GAGCAGAGAGCATCGTTCAGGCCAGGCC	CCAGCCACCTGTCTTAACTTGAAGGCA	2220
QY	2221	TTTTTGGGTCTCACGTGTCCACCCAGGG	GGGTGTCGGACTTTTGAACGGCTCTTACT	2280
Db	2221	TTTTTGGGTCTCACGTGTCCACCCAGGG	GGGTGTCGGACTTTTGAACGGCTCTTACT	2280
QY	2281	AAGAACGGGATGGGGTGGGGGGCTT	AGTGGGCTCTGGCTCACCTACAACTG	2340
Db	2281	AAGAACGGGATGGGGTGGGGGGCTT	AGTGGGCTCTGGCTCACCTACAACTG	2340
QY	2341	TGGTCAATGGGGTTATTTTAAACCCAG	GGGAAGAGGTATTTTGTTCACACAG	2400
Db	2341	TGGTCAATGGGGTTATTTTAAACCCAG	GGGAAGAGGTATTTTGTTCACACAG	2400
QY	2401	GGCCAGCAGGCTCTTGAATTTCGACCT	TTCGCAATTCGGGTAATTCGGTAA	2460

```
Db 2401 GGCAGGAGGCTCTTGAATTCGACCCCTTCGAGCTTGGCAATTCGGTACTGTTGGTAA 2460
Qy 2461 ATGAAGACGCCAAATCAATAAGAAAGGCGCGCCATTCCTATCTCTAGAGATGGA 2520
Db 2461 ATGAAGACGCCAAATCAATAAGAAAGGCGCGCCATTCCTATCTCTAGAGATGGA 2520
Qy 2521 ACCGCTGGAGGCAACTGCATTAAGGCTAGAGATACGCCCTTCTCTGGAACAATT 2580
Db 2521 ACCGCTGGAGGCAACTGCATTAAGGCTATGAAGATACGCCCTTCTCTGGAACAATT 2580
Qy 2581 GCTTTTACAGATGCACATATCGAGTGAACATCAGCTTCGCGGAATCTCGAAATGTC 2640
Db 2581 GCTTTTACAGATGCACATATCGAGTGAACATCAGCTTCGCGGAATCTCGAAATGTC 2640
Qy 2641 GTTTCGGTTGGCAGAAGCTATGAACGATATGGCTGAATACAAATCACAGAATCGTCGT 2700
Db 2641 GTTTCGGTTGGCAGAAGCTATGAACGATATGGCTGAATACAAATCACAGAATCGTCGT 2700
Qy 2701 ATGCAAGTAAATCTCTTCAATCTTTATGCGGTGTTGGGCCCTTTATTTATCCGGA 2760
Db 2701 ATGCAAGTAAATCTCTTCAATCTTTATGCGGTGTTGGGCCCTTTATTTATCCGGA 2760
Qy 2761 GTTGCAGTTCGCCGCCCGCGGAACA 2784
Db 2761 GTTGCAGTTCGCCGCCCGCGGAACA 2784
```

```
RESULT 2
AC105794
LOCUS Rattus norvegicus clone CH230-254E24, WORKING DRAFT SEQUENCE.
AC105794
VERSION AC105794.4 GI:24635611
HTG: HTGS PHASE2; HTGS DRAFT; HTGS_FULLTOP.
Rattus norvegicus (Norway rat)
SOURCE Rattus norvegicus
ORGANISM Rattus
```

```
REFERENCE
AUTHORS Muzny,D,Marie., Metzker,M, Lee., Abramson,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooke,S., Amin,A., Angulano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Blawalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cres,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falle,I., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Haves,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpachy,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuewa,L., Louised,H., Lozada,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S.,
Nwaokemele,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
```

FEATURES

Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C., Popper,F., Poidexter,A., Popovic,D., Primus,E., Pu,L.-L., Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R., Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F., Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J., Sanders,W., Savery,G., Scheier,S., Scott,G., Shatsman,S., Shen,H., Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajd,D., Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J., Steinle,M., Strong,R., Sutton,A., Svatek,A., Taber,P., Taylor,C., Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K., Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J., Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F., Williams,G., Willson,K., Wlarczyk,R., Wooden,H., Worley,K., Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V., Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von Niederhauser,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O., Weinstock,G. and Gibbs,R.A.

Direct Submission
Unpublished
2 (bases 1 to 311583)
Worley,K.C.
Direct Submission
Submitted (10-JAN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 311583)
Rat Genome Sequencing Consortium.
Direct Submission
Submitted (15-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA
On Nov 6, 2002 this sequence version replaced gi:23603171.
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GNRU
Center clone name: CH230-254E24
----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 203035 bases at least Q40
Consensus quality: 206075 bases at least Q30
Consensus quality: 207296 bases at least Q20
Estimated insert size: 209169; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 1 contigs. Gaps between the contigs are represented as runs of N. The order of the pieces is believed to be correct as given, however the sizes of the gaps between them are based on estimates that have been provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and the accession number will be preserved.
* 1 311583: contig of 311583 bp in length.
* Location/Qualifiers

[illegible]

[illegible]

Db 1141 CAATTGAGCAAGGATGGGGTGTGCAGAGGCTGCCCAAGATGCATAGGTGCAAAAGGCC 1200
QY 1642 TGAGGTTTCAGAGATGCTGTGATCCGGAATCAAAAGCTCAGGCTCTCCCTCTCTCTCTCTCC 1701
Db 1201 TGAGGTTTCAGAGATGCTGTGATCCGGAATCAAAAGCTCAGGCTCTCCCTCTCTCTCTCC 1260
QY 1702 TCCTCTGCCGCCCTCCCTCTCTCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 1761
Db 1261 TCCTCTGCCGCCCTCCCTCTCTCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 1320
QY 1762 TCCT 1821
Db 1321 TCCT 1380
QY 1822 CTTCTGCCGCCCT 1881
Db 1381 CTTCTGCCGCCCT 1440
QY 1882 TACCT 1941
Db 1441 TACCT 1500
QY 1942 CCT 2001
Db 1501 CCT 1560
QY 2002 TCCCTCCCT 2061
Db 1561 TCCCTCCCT 1620
QY 2062 CTTCT 2121
Db 1621 CTTCT 1680
QY 2122 CTTGGGTTACTTTTCCCAATTAGCAATGGCAGGACCCAGAGACACAGAGATCGTTTCCCA 2181
Db 1681 CTTGGGTTACTTTTCCCAATTAGCAATGGCAGGACCCAGAGACACAGAGATCGTTTCCCA 1740
QY 2182 GGCAGGCCCCAGCACTGTCTTTAACTTGAAGGCAATTTTGGGTCTCAGGTGTCCA 2241
Db 1741 GGCAGGCCCCAGCACTGTCTTTAACTTGAAGGCAATTTTGGGTCTCAGGTGTCCA 1800
QY 2242 CCAGGCGGGTGTGCGACTTTGAACGGCTCTTACTTTCAGAAAGCGCATGGGGTGGGG 2301
Db 1801 CCAGGCGGGTGTGCGACTTTGAACGGCTCTTACTTTCAGAAAGCGCATGGGGTGGGG 1860
QY 2302 GGCTTAGGTGGCTCTGCTCTCACTACAAGTGCACAAAGTGTGTCATGGGGTTATTTTAA 2361
Db 1861 GGCTTAGGTGGCTCTGCTCTCACTACAAGTGTGTCATGGGGTTATTTTAA 1920
QY 2362 CCCAGGGAAGGATTTATTTGTTCCAGAGGCGGGCGGCGAGAGCTCTCTTGAAT 2421
Db 1921 CCCAGGGAAGGATTTATTTGTTCCAGAGGCGGGCGGCGAGAGCTCTCTTGAAT 1980

RESULT 4

LOCUS S80994 1964 bp DNA linear ROD 10-MAY-1996
DEFINITION MLC-2-myosin light chain 2 (promoter) [rats, Wistar-Kyoto,
spontaneously hypertensive, Genomic, 1964 nt].
ACCESSION S80994
VERSION S80994.1 GI:1311693
KEYWORDS Rattus sp.
SOURCE Rattus sp.
ORGANISM Rattus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (bases 1 to 1964)
Doud,S.K., Pan,L.X., Carleton,S., Marmorstein,S. and Siddiqui,M.A.
AUTHORS Adaptational response in transcription factors during development
TITLE of myocardial hypertrophy
J. Mol. Cell. Cardiol. 27 (10), 2359-2372 (1995)
JOURNAL

MEDLINE 96129594
PUBMED 8576950
REMARK GenBank staff at the National Library of Medicine created this
entry [NCBI gisbseq 175713] from the original journal article.
FEATURES
Location/Qualifiers
1..1964
/organism="Rattus sp."
/mol_type="genomic DNA"
/db_xref="taxon:10118"
1..1964
/gene="myosin light chain 2"
/note="myosin light chain 2"

ORIGIN

Query Match 70.4%; Score 1959.2; DB 10; Length 1964;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1961; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 442 GCCCATGTGGCTGGGCGCCGATGTTTCCAGAAATCCTCTGGTCTAAGGATCCAGACCTCTT 501
Db 1 GCCCATGTGGCTGGGCGCCGATGTTTCCAGAAATCCTCTGGTCTAAGGATCCAGACCTCTT 60
QY 502 ACGGAGCCCAACAGCTCAAGGACAGATTAGCATGTTTCTATGTGTACTGGGAGGAGGAG 561
Db 61 ACGGAGCCCAACAGCTCAAGGACAGATTAGCATGTTTCTATGTGTACTGGGAGGAGGAG 120
QY 562 CCAACAGAGGTCATCAAGATCCACAGGGGCTCCGGTTCGGAGGCCCTTGGGTTTATCAC 621
Db 121 CCAACAGAGGTCATCAAGATCCACAGGGGCTCCGGTTCGGAGGCCCTTGGGTTTATCAC 180
QY 622 CAAATGTTTCCACCAGCAACATAAAACAGCTCCTCAGACAGCGCAGTGGACCACTGGA 681
Db 181 CAAATGTTTCCACCAGCAACATAAAACAGCTCCTCAGACAGCGCAGTGGACCACTGGA 240
QY 682 CCACAGGACAGATCACCTCTGTGGGCCAGACTCATAGTAACCTCTCACTCAATCTCC 741
Db 241 CCACAGGACAGATCACCTCTGTGGGCCAGACTCATAGTAACCTCTCACTCAATCTCC 300
QY 742 AGCCTCCACAGTCAATGTGGTCACTTGTTCACGACCACACACTTGGCAAGTCAG 801
Db 301 AGCCTCCACAGTCAATGTGGTCACTTGTTCACGACCACACACTTGGCAAGTCAG 360
QY 802 TGTGCTCAACAACATCTTTCAGAAAGCGGGGATGGGTTTGTAACTGATGGGTGT 861
Db 361 TGTGCTCAACAACATCTTTCAGAAAGCGGGGATGGGTTTGTAACTGATGGGTGT 420
QY 862 TTTGTTTGTGTTTGTGTTTCAATTAAGTGTGTCAGTGGGCCAGGCTAGGCTTGAACCTCATG 921
Db 421 TTTGTTTGTGTTTGTGTTTCAATTAAGTGTGTCAGTGGGCCAGGCTAGGCTTGAACCTCATG 480
QY 922 TAGGCAAGCATGACCATGAACCTCTCTGATCCTCTCTCAGTGTCTGGGATAACAGGTG 981
Db 481 TAGGCAAGCATGACCATGAACCTCTCTGATCCTCTCTCAGTGTCTGGGATAACAGGTG 540
QY 982 TGTGTCACTCCCTACCTCTTAATAGCAATATGTGGCCACATGTTTGTGCCCCACAGGTT 1041
Db 541 TGTGTCACTCCCTACCTCTTAATAGCAATATGTGGCCACATGTTTGTGCCCCACAGGTT 600
QY 1042 GAGACATCTTGACCTGAGGAAGAAATAGCTAACTCACTCACTCTGAGGTTGCTCGAT 1101
Db 601 GAGACATCTTGACCTGAGCAACAAATAGTAACTCACTCACTCTGAGGTTGCTCGAT 660
QY 1102 CTCGTCTTGTCTTTCAGACACTCAGGAGTGGGGGTGAGAAAGTCAAGTCAAGCCCT 1161
Db 661 CTCGTCTTGTCTTTCAGACACTCAGGAGTGGGGGTGAGAAAGTCAAGTCAAGCCCT 720
QY 1162 GCTACATAATAGTTCAAGGCTCGCTGAGTACATGAGACCATGCTCTCAAAAGAAAG 1221
Db 721 GCTACATAATAGTTCAAGGCTCGCTGAGTACATGAGACCATGCTCTCAAAAGAAAG 780
QY 1222 GAATTGGTATAGTACATCTCTGCTCTCCAGTACTTGGGACACAGAGGCCACTCCA 1281
Db 781 GAATTGGTATAGTACATCTCTGCTCTCCAGTACTTGGGACACAGAGGCCACTCCA 840

QY	1282	CCACCATCTCCAGCAGCTGGGCTGCCTCCCGAGCCTCGTTATTTCATATCAATGAGAT	1341
DB	841		
		CCAACAATCTCCAGCAGCTGGGCTGCCTCCCGAGCCTCGTTATTTCATATCAATGAGAT	900
QY	1342	GGGACCCAACTGCTAAAGGTGACCTTGCACCCACGGGTGACTTGGAGACTTGGAGGTGGA	1401
DB	901	GGGACCCAACTGCTAAAGTGACCTTGCACCCACGGGTGACTTGGAGACTTGGAGGTGGA	960
		GGGTTTATCATTTCTCCAGTCGGTCAGCAAGTGCTCGCCGCAAGAAGGTTTTGAGTTCA	1020
QY	1402	GGGTTTATCATTTCTCCAGTCGGTCAGCAAGTGCTCGCCGCAAGAAGGTTTTGAGTTCA	1461
DB	961	GGGTTTATCATTTCTCCAGTCGGTCAGCAAGTGCTCGCCGCAAGAAGGTTTTGAGTTCA	1020
		AAGTAGAAGATGGACACAGGAGACACAGGAGAGACCCACCTCGAGCTGACTGTCTC	1521
QY	1462	AAGTAGAAGATGGACACAGGAGACACAGGAGAGACCCACCTCGAGCTGACTGTCTC	1521
DB	1021		
		AAGTAGAAGATGGACACAGGAGAGACACAGGAGAGACCCACCTCGAGCTGACTGTCTC	1080
QY	1522	CTGTGCGGCTGGGTGGGGACACAAGACAGAGAAGCAGAGGCAGAGAAACAAGGGTGGGTGA	1581
DB	1081	CTGTGCGGCTGGGTGGGGACACAAGACAGAGAAGCAGAGGCAGAGAAACAAGGGTGGGTGA	1140
		CATTGTAGCAAGAGATGGGGTGTGCCAGAGGCTGCCCAAGATGATAGGTGCAAAAGGCC	1641
QY	1582	CATTGTAGCAAGAGATGGGGTGTGCCAGAGGCTGCCCAAGATGATAGGTGCAAAAGGCC	1641
DB	1141	CATTGTAGCAAGAGATGGGGTGTGCCAGAGGCTGCCCAAGATGATAGGTGCAAAAGGCC	1200
		TGAGGTTGAGGATGCCTGGATCCGGAATCAAAAGCTCAGGCTCCTCCCTCTTCTCTCTCC	1701
QY	1642	TGAGGTTGAGGATGCCTGGATCCGGAATCAAAAGCTCAGGCTCCTCCCTCTTCTCTCTCC	1701
DB	1201	TGAGGTTGAGGATGCCTGGATCCGGAATCAAAAGCTCAGGCTCCTCCCTCTTCTCTCTCC	1260
		TCCTCTGCCCCCTCCTCCTCTCTGCCCCCTCTTCTCTCTGCCCCCTCTTCTCTCTCC	1761
QY	1702	TCCTCTGCCCCCTCCTCCTCTCTGCCCCCTCTTCTCTCTGCCCCCTCTTCTCTCTCC	1761
DB	1261	TCCTCTGCCCCCTCCTCCTCTCTGCCCCCTCTTCTCTCTGCCCCCTCTTCTCTCTCC	1320
		TCCTCTTCTCTCCTCCTCCCTCTCATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
QY	1762	TCCTCTTCTCTCCTCCTCCTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1821
DB	1321	TCCTCTTCTCTCCTCCTCCCTCTCATCTACTCTCTCTCTCTCTCTCTCTCTCTCTCT	1380
		CCTCTGCCCCCTCTTCT	1881
QY	1822	CCTCTGCCCCCTCTTCT	1881
DB	1381		
		CCTCTGCCCCCTCTTCT	1440
QY	1882	TACCTCTCTTCT	1941
DB	1441	TACCTCTCTTCT	1500
		CCCTCTTCT	2001
QY	1942	CCCTCTTCT	2001
DB	1501	CCCTCTTCT	1560
		TCCTCCCTCTCTCATCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2061
QY	2002	TCCTCCCTCTCTCATCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	2061
DB	1561	TCCTCCCTCTCTCATCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1620
		CCTCCCT	2121
QY	2062	CCTCCCT	2121
DB	1621	CCTCCCT	1680
		CCTGGGTTACTTTTCCCCCAATAGACAATGGCAGACCCAGAGCACAGAGCATCGTTCCCA	2181
QY	2122	CCTGGGTTACTTTTCCCCCAATAGACAATGGCAGACCCAGAGCACAGAGCATCGTTCCCA	2181
DB	1681	CCTGGGTTACTTTTCCCCCAATAGACAATGGCAGACCCAGAGCACAGAGCATCGTTCCCA	1740
		GGCCAGGCCCCAGGCACCTGTCTTTTAACTTGAAGGCATTTTGGGTCCTACAGTGCCA	2241
QY	2182	GGCCAGGCCCCAGGCACCTGTCTTTTAACTTGAAGGCATTTTGGGTCCTACAGTGCCA	2241
DB	1741	GGCCAGGCCCCAGGCACCTGTCTTTTAACTTGAAGGCATTTTGGGTCCTACAGTGCCA	1800
		CCCAGGCGGTGTCCGACTTTTGAACGGCTCTTACTTTCAGAGAAGCGCATGGGTGGGG	2301
QY	2242	CCCAGGCGGTGTCCGACTTTTGAACGGCTCTTACTTTCAGAGAAGCGCATGGGTGGGG	2301
DB	1801	CCCAGGCGGTGTCCGACTTTTGAACGGCTCTTACTTTCAGAGAAGCGCATGGGTGGGG	1860
		GGCTTAGTGGCCTCTGCCTCACCTACAACTGCCAAAAGTGGTCATGGGGTATTTTAA	2361
QY	2302	GGCTTAGTGGCCTCTGCCTCACCTACAACTGCCAAAAGTGGTCATGGGGTATTTTAA	2361
DB	1861	GGCTTAGTGGCCTCTGCCTCACCTACAACTGCCAAAAGTGGTCATGGGGTATTTTAA	1920

Qy	2362	CC	CCAGGAGAGAGGTATTTATGTTCCACAGACAGGGCGCGCCA	2405
Db	1921	CC	CCAGGAGAGAGGTATTTATGTTCCACAGACAGGGCGCGCCA	1964
RESULT 5	AC118330_2/c			
WPCOMMENT				
		Sequence split into 5 fragments	LOCUS AC118330	Accession AC118330
		Fragment Name	Begin	End
		AC118330_0	1	110000
		AC118330_1	100001	210000
		AC118330_2	200001	310000
		AC118330_3	300001	410000
		AC118330_4	400001	509911
		Continuation (3 of 5) of AC118330	from base 200001	(AC118330 Rattus norvegicus cl
		Query Match	52.9%	Score 1473.6; DB 2; Length 110000;
		Best Local Similarity	66.9%;	Pred. No. 0;
		Matches 1515;	Conservative	0; Mismatches 719; Indels 29; Gaps 1;
Qy	189	GGTACCCAGGACTGATTCTCGGAAGTTCTAGGCTGCAGAAATCTCACGACCAAGAGT	248	
Db	106580	GGTACCCAGGACTGATTCTCGGAAGTTCTAGGCTGCAGAAATCTCACGACCAAGAGT	106521	
Qy	249	TTGGAGTCACAGGATGGGTCTCGCCAAGAGCCTAGGACAGAGAGTTGTCTAGCCCCCTGT	308	
Db	106520	TTGGAGTCACAGGATGGGTCTCGCCAAGAGCCTAGGACAGAGAGTTGTCTAGCCCCCTGT	106461	
Qy	309	CCCCGGACCCCTGTGGACTGTGAGAAGAGCAGAGTCCCCACCCCGAGGCTTCTTAGACCCA	368	
Db	106460	CCCCGGACCCCTGTGGACTGTGAGAAGAGCAGAGTCCCCACCCCGAGGCTTCTTAGACCCA	106401	
Qy	369	CCCCGGGTTTTCCAGCATCTTCTCTGAGACCGGACCCCTGGCTGAAGATACAGAAAC	428	
Db	106400	CCCCGGGTTTTCCAGCATCTTCTCTGAGACCGGACCCCTGGCTGAAGATACAGAAAC	106341	
Qy	429	CCTAGAGTCTGACGCCATGTGGCTGGGCGCCATGTTTCCAGAACTCTCTGGTCTTAAGG	488	
Db	106340	CCTAGAGTCTGACGCCATGTGGCTGGGCGCCATGTTTCCAGAACTCTCTGGTCTTAAGG	106281	
Qy	489	ATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGACAGTTAGCATGTTTCATGTGTA	548	
Db	106280	ATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGACAGTTAGCATGTTTCATGTGTA	106221	
Qy	549	GGAGAGCAGAGGCCAACAGAGGTCTATGAAGATCCAGGGGCTCCGGTTCGAGGCCCT	608	
Db	106220	GGAGAGCAGAGGCCAACAGAGGTCTATGAAGATCCAGGGGCTCCGGTTCGAGGCCCT	106161	
Qy	609	TGGGTTTTATCACCMAATGTTTTCCACCCACGACATAAACAGCTCTCTAGACAGCGCA	668	
Db	106160	TGGGTTTTATCACCMAATGTTTTCCACCCACGACATAAACAGCTCTCTAGACAGCGCA	106101	
Qy	669	GTGGACCAGTGGACACAGGACAGATCACCTCTGTGGGCCACAGACTCATAGTAACCTCT	728	
Db	106100	GTGGACCAGTGGACACAGGACAGATCACCTCTGTGGGCCACAGACTCATAGTAACCTCT	106041	
Qy	729	AACCTCAATCTCCAGCTCCACAGTCTATTGTCGGTCACTTGTGTTTCTCAGCCACACAC	788	
Db	106040	AACCTCAATCTCCAGCTCCACAGTCTATTGTCGGTCACTTGTGTTTCTCAGCCACACAC	105981	
Qy	789	TTGCGAGTCACTGTGCTCAACACATCTTCAGAGCCAGGGGGATGGGGTTTGTGTT	848	
Db	105980	TTGCGAGTCACTGTGCTCAACACATCTTCAGAGCCAGGGGGATGGGGTTTGTGTT	105921	
Qy	849	AACGTATGGGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG	908	
Db	105920	AACGTATGGGTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTGTTTTG	105861	
Qy	909	TGAATCTACTATGATAGGACGATGACCATGAATCTTCTGATCTCTCTCTCAGTGTCT	968	
Db	105860	TGAATCTACTATGATAGGACGATGACCATGAATCTTCTGATCTCTCTCTCAGTGTCT	105801	

QY 969 GGATAACAGGTGTGTCTCACTCCCTACCCCTTCTAATAGCAATATGTGGCCACATGTTTG 1028
Db 105800 GGGATAACAGGTGTGTCTCACTCCCTACCCCTTCTAATAGCAATATGTGGCCACATGTTTG 105741
QY 1029 TGCCCAACAGGTTGAGACCACTCTTGACCTGAGGAAGAAATAGCTAACTACCTCCTGA 1088
Db 105740 TGCCCAACAGGTTGAGACCACTCTTGACCTGAGGAAGAAATAGCTAACTACCTCCTGA 105681
QY 1089 AGGTTCCCTGGATCTCGTCTTTGCTTTCCAGCACTCAGAGTGGGGGGTCTCAGAGTGC 1148
Db 105680 AGGTTCCCTGGATCTCGTCTTTGCTTTCCAGCACTCAGAGTGGGGGGTCTCAGAGTGC 105621
QY 1149 AAAGTCAGCCCTGCTACATAATAGATTCAAGGCTCGCCTGGGCTACATGAGACCATGCC 1208
Db 105620 AAAGTCAGCCCTGCTACATAATAGATTCAAGGCTCGCCTGGGCTACATGAGACCATGCC 105561
QY 1209 TCAAAAGAAAGAAATTTGGTATAGTGACATATCTTGTCCTCCAGTACTTAGGGACAC 1268
Db 105560 TCAAAAGAAAGAAATTTGGTATAGTGACATATCTTGTCCTCCAGTACTTAGGGACAC 105501
QY 1269 AGAGGCCACTCCACACCACTCTCCAGCAGCTGGCTCCCTCCCGAGCCTCGTTTATTTC 1328
Db 105500 AGAGGCCACTCCACACCACTCTCCAGCAGCTGGCTCCCTCCCGAGCCTCGTTTATTTC 105441
QY 1329 ATATCAATGAGATGGGACCC-----AACTGCTAAG 1359
Db 105440 ATATCAATGAGATGGGACCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 105381
QY 1360 GTGACCTTGCACCCACGGGTGACTGGAGCTTGAGAGTGGAGGTTTATCATTTCTCCA 1419
Db 105380 NNN 105321
QY 1420 GTCGGTCAGCAAGTGTGTCGCCCAAGAGGTTTTCAGTTCAAGTAGAAGATGGACAG 1479
Db 105320 NNN 105261
QY 1480 GGAGAGACCAAG 1539
Db 105260 NNN 105201
QY 1540 ACACAAAGCAG 1599
Db 105200 NNN 105141
QY 1600 GGTGTGCCAGAGGCTGCCAAGATGCTAGGTGCAAGGCCCTTGAGGTTCGAGGATGCCT 1659
Db 105140 NNN 105081
QY 1660 GGATCCGGAATCAAAGCTCAGGCTCCTCCCTCTCTCCTCTCTCTCTCTCTCTCTCT 1719
Db 105080 NNN 105021
QY 1720 TCCTCTGCCCCCT 1779
Db 105020 NNN 104961
QY 1780 CCTCATCTACT 1839
Db 104960 NNN 104901
QY 1840 CCT 1899
Db 104900 NNN 104841
QY 1900 CTCCT 1959
Db 104840 NNN 104781
QY 1960 TCTTCT 2019
Db 104780 NNN 104721
QY 2020 CTTCT 2079

Db 104720 NNN 104661
QY 2080 CCT 2139
Db 104660 CCT 104601
QY 2140 ATTAGACAATGCGAG 2199
Db 104600 ATTAGACAATGCGAG 104541
QY 2200 GTCTCTTTAACTTTGAAGGCAATTTTGGGTCTCAGCTGTCCACCCAGGCGGTCTCGGAC 2259
Db 104540 GTCTCTTTAACTTTGAAGGCAATTTTGGGTCTCAGCTGTCCACCCAGGCGGTCTCGGAC 104481
QY 2260 TTTGAACGGCTCTTACTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2319
Db 104480 TTTGAACGGCTCTTACTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 104421
QY 2320 CTCACCTCAACTGCAAAAGTGTGTCATGGGTATTTTAACTCCAGGAGAGAGAGATTT 2379
Db 104420 CTCACCTCAACTGCAAAAGTGTGTCATGGGTATTTTAACTCCAGGAGAGAGAGATTT 104361
QY 2380 TATTGTTCCACAGCAGGCGCGCCAGCAGGCTCTTGAAATTC 2422
Db 104360 TATTGTTCCACAGCAGGCGCGCCAGCAGGCTCTTGAAATTC 104318

RESULT 6
AC118330.3/c
WPCOMMENT
Sequence split into 5 fragments LOCUS AC118330 Accession AC118330
Fragment Name Begin End
AC118330_0 1 110000
AC118330_1 100001 210000
AC118330_2 200001 310000
AC118330_3 300001 410000
AC118330_4 400001 509911
Continuation (4 of 5) of AC118330 from base 300001 (AC118330 Rattus norvegicus clone CH2;
Query Match 52.9%; Score 1473.6; DB 2; Length 110000;
Best Local Similarity 66.9%; Pred. No. 0;
Matches 1515; Conservative 0; Mismatches 719; Indels 29; Gaps 1;

QY 189 GGTACCCAGGACTGATTCTCGGAAGTTCTAGGCTGCAGAAATCTCAGACGCACAGAGT 248
Db 6580 GGTACCCAGGACTGATTCTCGGAAGTTCTAGGCTGCAGAAATCTCAGACGCACAGAGT 6521
QY 249 TTGGAGTCACAGATGGGTGCTCGCAAGAGCTTAGGGACAGAGCTTGTTCAGCCCCCTGT 308
Db 6520 TTGGAGTCACAGATGGGTGCTCGCAAGAGCTTAGGGACAGAGCTTGTTCAGCCCCCTGT 6461
QY 309 GCCCGAGCCCTGTGACTGTGTAGAAAGAGAGAGTCCACCCCGAGGCTTCTTAGACCCA 368
Db 6460 GCCCGAGCCCTGTGACTGTGTAGAAAGAGAGTCCACCCCGAGGCTTCTTAGACCCA 6401
QY 369 CCCCGGTTTTCCAGCATCTTCTCGAGGACCGAGCCCTCGGTGAAAGTACAGAAAC 428
Db 6400 CCCCGGTTTTCCAGCATCTTCTCGAGGACCGAGCCCTCGGTGAAAGTACAGAAAC 6341
QY 429 CCTAGAGTCGAGGCCCATGTGGCTGGCGCCATGTTTCCAGAAATCTCTGTCTTAAGG 488
Db 6340 CCTAGAGTCGAGGCCCATGTGGCTGGCGCCATGTTTCCAGAAATCTCTGTCTTAAGG 6281
QY 489 ATCCAGACTCTTACGGAGCCCAACAGCTCAAGGAGAGTTAGCATGTTTCTGTGTACTG 548
Db 6280 ATCCAGACTCTTACGGAGCCCAACAGCTCAAGGAGAGTTAGCATGTTTCTGTGTACTG 6221
QY 549 GGAGGAGCAGGAGCCCAACAGAGGTCTATGAAGATCCACAGGGGCTCCGGTTCGAGGCCCT 608
Db 6220 GGAGGAGCAGGAGCCCAACAGAGGTCTATGAAGATCCACAGGGGCTCCGGTTCGAGGCCCT 6161
QY 609 TGGGTTTTATCACCAATGTTTTCCACCAGCAACATAAACAGCTCTCTCAGACGCGCA 668

Db 6160 TGGGTTTTATACAAATGTTTCCACCCAGCAACATAAAACAGCTCCTCAGACAGCGCA 6101
Qy 669 GTGACCAAGTGGACCAAGGACAGATCACTCTGTGGGCCAGACTCATAGTAACCTCT 728
Db 6100 GTGACCAAGTGGACCAAGGACAGATCACTCTGTGGGCCAGACTCATAGTAACCTCT 6041
Qy 729 AACCTCAATCTCAGCCTCCACAGTCAATGTGGGTCACTTGTTCATGACCAACACAC 788
Db 6040 AACCTCAATCTCAGCCTCCACAGTCAATGTGGGTCACTTGTTCATGACCAACACAC 5981
Qy 789 TTGSCAAGTCAGGTGSCCTCAACACAATCTCAGAGCCAGGGGATGGGTTTGT 848
Db 5980 TTGSCAAGTCAGGTGSCCTCAACACAATCTCAGAGCCAGGGGATGGGTTTGT 5921
Qy 849 AACCTGATGGGTGTTTGTGTTTGTTCATTAACCTGACGTAGCCAGGCTAGCCT 908
Db 5920 AACCTGATGGGTGTTTGTGTTTGTTCATTAACCTGACGTAGCCAGGCTAGCCT 5861
Qy 909 TGAACCTCACTATGAGCAAGCATGACCATGAACCTCTGATCCTCTCTCATGTGCT 968
Db 5860 TGAACCTCACTATGAGCAAGCATGACCATGAACCTCTGATCCTCTCTCATGTGCT 5801
Qy 969 GGGATAACAGGTGTGTCACTCCCTACCTCTTAATAGCAATATGTGGCCACATGTTG 1028
Db 5800 GGGATAACAGGTGTGTCACTCCCTACCTCTTAATAGCAATATGTGGCCACATGTTG 5741
Qy 1029 TGCCCCACAGGTTGAGACCATCTTGACCTGAGGAAGAAATAGCTAACACTCACCTCTGA 1088
Db 5740 TGCCCCACAGGTTGAGACCATCTTGACCTGAGGAAGAAATAGCTAACACTCACCTCTGA 5681
Qy 1089 AGGTGCTCGATCTCGTCTTTGTCTTTCAGCACTCAGAGTGGGGGTGCAAGATGC 1148
Db 5680 AGGTGCTCGATCTCGTCTTTGTCTTTCAGCACTCAGAGTGGGGGTGCAAGATGC 5621
Qy 1149 AAGTCAGCCCTGCTACATAATGAGTTCAGGCTCGCCTGGGCTACATGAGCCATGCC 1208
Db 5620 AAGTCAGCCCTGCTACATAATGAGTTCAGGCTCGCCTGGGCTACATGAGCCATGCC 5561
Qy 1209 TCAAAAGAAAGAAATGGTATAGTACATATCTGGTCTCCAGTACTTAGGGACAC 1268
Db 5560 TCAAAAGAAAGAAATGGTATAGTACATATCTGGTCTCCAGTACTTAGGGACAC 5501
Qy 1269 AGAGGCCACTCCACCACTCTCCAGCAGTGGCTCCCTCCCGAGCCTCGTTATTTC 1328
Db 5500 AGAGGCCACTCCACCACTCTCCAGCAGTGGCTCCCTCCCGAGCCTCGTTATTTC 5441
Qy 1329 ATATCAATGATGGGACCC-----AAGTCTAAG 1359
Db 5440 ATATCAATGATGGGACCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 5381
Qy 1360 GTGACCTTGCAACCACGGGGTGAAGTGGAGACTTGAGAGTGGAGGTTTATCATTTCTCCA 1419
Db 5380 NNN 5321
Qy 1420 GTCGGTCAGCAAGTGGTCGCCGCCAAGAGGTTTGTAGTTCAAAGTAGAAGTGGACAG 1479
Db 5320 NNN 5261
Qy 1480 GGAGAGACCGAGGAAGACCCACCTGGAGCTGACTGCTCGCGCTGGTGGGG 1539
Db 5260 NNN 5201
Qy 1540 ACACAAAGCAGAGAAGCAGAGGCAGAGAACAGGGTGGTGACATTTGAGCAAGGATGGG 1599
Db 5200 NNN 5141
Qy 1600 GGTGTGCGCAGAGGCTGCCAAGATGCATAGTGTCAAGGCCCTGAGGTTCCAGGATGCCT 1659
Db 5140 NNN 5081
Qy 1660 GGATCCGGAATCAAGCTCAGGCTCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCT 1719

Db 5080 NNN 5021
Qy 1720 TCCTCTGCCCT 1779
Db 5020 NNN 4961
Qy 1780 CCTCATCTACT 1839
Db 4960 NNN 4901
Qy 1840 CCT 1899
Db 4900 NNN 4841
Qy 1900 CTCCT 1959
Db 4840 NNN 4781
Qy 1960 TCTTCT 2019
Db 4780 NNN 4721
Qy 2020 CCT 2079
Db 4720 NNN 4661
Qy 2080 CTTCT 2139
Db 4660 CTTCT 4601
Qy 2140 ATTAGACAATGGCAGACCCAGAGCAGACAGACATCGTTCAGGCGCAGGCCAGCCACT 2199
Db 4600 ATTAGACAATGGCAGACCCAGAGCAGACAGACATCGTTCAGGCGCAGGCCAGCCACT 4541
Qy 2200 GTCTCTTTAACTTGAAGCATTTTGGGTCTCACGTGTCCACCCAGGCGGGTGGGAC 2259
Db 4540 GTCTCTTTAACTTGAAGCATTTTGGGTCTCACGTGTCCACCCAGGCGGGTGGGAC 4481
Qy 2260 TTTGAACGGCTCTTACTTTCAGAAAGCGCATGGGGTGGGGGGCTTAGGTGGCTCTGC 2319
Db 4480 TTTGAACGGCTCTTACTTTCAGAAAGCGCATGGGGTGGGGGGCTTAGGTGGCTCTGC 4421
Qy 2320 CTCACCTACAATGCCAAAGTGGTCATGGGGTATTTTAAACCAGGGAAGAGGTATT 2379
Db 4420 CTCACCTACAATGCCAAAGTGGTCATGGGGTATTTTAAACCAGGGAAGAGGTATT 4361
Qy 2380 TATTGTTCCACAGCAGCGCGCCAGCAGGCTCTTGAATTC 2422
Db 4360 TATTGTTCCACAGCAGCGCGCCAGCAGGCTCTTGAATTC 4318

RESULT 7

AC137306 237801 bp DNA linear HTG 20-NOV-2002
LOCUS Rattus norvegicus clone CH230-unknown, *** SEQUENCING IN PROGRESS
DEFINITION *** 5 unordered pieces.

AC137306

AC137306.1 GI:25138383

HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_ENRICHED.

Rattus norvegicus (Norway rat)

KEYWORDS

Rattus norvegicus

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

ORGANISM

Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE

1 (bases 1 to 237801)

AUTHORS

Muzny, D., Marie, Metzker, M., Lee, A., Adams, C., Alder, J., Allen, C., Allen, H., Alsbrooks, S., Amin, A., Anguiano, D., Anyalebechi, V., Ayoyagi, A., Ayodeji, M., Baca, E., Baden, H., Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benahmed, F., Biswal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M., Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, E., Cardenas, V., Carter, K., Cavazos, I., Cesar, H., Center, A., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chu, J.,

Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L., Davila, M.L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D., Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Divya, K., Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escoto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Frazer, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregiorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamil, C., Hamilton, C., Hamilton, K., Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, S., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Huylk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Kraft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensuhow, L., Louleghed, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindartne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhiney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montanayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munidasa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwaokelimeh, O., Okwuonu, G., Olarnpunsagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.-L., Puato, M., Quiroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartsbeyn, A., Sisson, I., Sitter, C.D., Smajs, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steinle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, F., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Usmani, K., Valas, R., Vera, V., Villaseana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F., Williams, G., Willson, R., Wleczky, R., Wooden, H., Worley, K., Yu, F., Zhang, J., Zhou, X., Zhou, X., Zhao, S., Yoon, L., Yoon, V., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Niederhauser, J., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstein, G., and Gibbs, R.A.

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

2 (bases 1 to 237801)
Rat Genome Sequencing Consortium.

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu

----- Project Information
Center project name: K2UP
Center clone name: CH230-unknown

----- Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 191019 bases at least Q40
Consensus quality: 195314 bases at least Q30

Consensus quality: 198237 bases at least Q20
Estimated insert size: 194634; sum-of-contigs estimation
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 5 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 3881: contig of 3881 bp in length
* 3882 3981: gap of unknown length
* 3982 230605: contig of 226624 bp in length
* 230606 230705: gap of unknown length
* 230706 231774: contig of 1069 bp in length
* 231775 231874: gap of unknown length
* 231875 232328: contig of 1364 bp in length
* 232329 233338: gap of unknown length
* 233339 237801: contig of 4463 bp in length.

FEATURES
source

1..237801
/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-unknown"

1..1145

misc_feature

/note="wgs_contig"

misc_feature

3982..5025
/note="wgs_contig"

misc_feature

75038..77999
/note="wgs_contig"

ORIGIN

Query Match 48.1%; Score 1338; DB 2; Length 237801;
Best Local Similarity 79.0%; Pred. No. 1.1e-310;
Matches 1765; Conservative 0; Mismatches 55; Indels 414; Gaps 1;
QY 189 GGTACCCAGGACTGATTCTCGGAAAGTTCTAGGCTGCAGAAATCTCACAGCACAAGAGT 248
Db 131424 GGTACCCAGGACTGATTCTCGGAAAGTTCTAGGCTGCAGAAATCTCACAGCACAAGAGT 131483
QY 249 TTGGAGTCACAGGATGGGTGCTCGGCAAGAGCTTAGGACAGAGCGTTGTTCAGCCCTGT 308
Db 131484 TTGGAGTCACAGGATGGGTGCTCGGCAAGAGCTTAGGACAGAGCGTTGTTCAGCCCTGT 131543
QY 309 GCCCGGACCTGTGTGACTGTGAGAGAGCAGAGTCCACCCAGGCGCTTCTTTAGACCCA 368
Db 131544 GCCCGGACCTGTGTGACTGTGAGAGAGCAGAGTCCACCCAGGCGCTTCTTTAGACCCA 131603
QY 369 CCCCGGGTTTTCCAGCATCTCTTCGAGGACCGGACCCCTGGCTGAAAGTACAGAAAC 428
Db 131604 CCCCGGGTTTTCCAGCATCTCTTCGAGGACCGGACCCCTGGCTGAAAGTACAGAAAC 131663
QY 429 CCTAGAGTCGAGCCCATGTGGCTGGCGCGCATGTTTCAGAAATCTCTCGTCTAAGG 488
Db 131664 CCTAGAGTCGAGCCCATGTGGCTGGCGCGCATGTTTCAGAAATCTCTCGTCTAAGG 131723
QY 489 ATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGACAGTGTAGCATGTTCTGTGTACTG 548
Db 131724 ATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGACAGTGTAGCATGTTCTGTGTACTG 131783
QY 549 GGAGGACGAGGAGCCCAACAGAGGTTCATGAAGATCCACAGGGGCTCCGGTTCGAGGCCCT 608
Db 131784 GGAGGACGAGGAGCCCAACAGAGGTTCATGAAGATCCACAGGGGCTCCGGTTCGAGGCCCT 131843
QY 609 TGGGTTTTATACCAAAATGTTTCCACCAGCAACATAAACAGCTCTCTCAGACAGCGCA 668
Db 131844 TGGGTTTTATACCAAAATGTTTCCACCAGCAACATAAACAGCTCTCTCAGACAGCGCA 131903

[illegible][illegible]

PECIIT. 8

REV 8
AC137290/C

LOCUS 10631CTCY

DEFINITION

5444

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISMS

REFERENCE

AUTHORS

Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K., Egan, A., Escotto, M., Eugene, C., Evans, C.A., Falls, T., Fan, G., Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P., Fraser, C.M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M., Gebregeorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W., Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K., Harvey, Y., Haylak, P., Hawes, A., Henderson, N., Hernandez, J., Hernandez, R., Hines, S., Hladun, S.L., Hodgson, A., Hogues, M., Hollins, B., Howells, S., Hulyk, S., Hume, J., Idlebird, D., Jackson, A., Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A., Karpathy, S., Kelly, S., Kelly, S., Khan, Z., King, L., Kovar, C., Kowis, C., Krafft, C.L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J., Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J., Lorensu, H., Louleseg, H., Lozano, R.J., Lu, X., Ma, J., Maheshwari, M., Mahindaratne, M., Mahmoud, M., Malloy, K., Mangum, A., Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E., Mawhney, S., McLeod, M.P., McNeill, T.Z., Meenen, E., Milosavljevic, A., Miner, G., Minja, E., Montemayor, J., Moore, S., Morgan, M., Morris, K., Morris, S., Munitadisa, M., Murphy, M., Nair, L., Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S., Nwako, O., Okwuonu, G., Olarnpunagoon, A., Pal, S., Parks, K., Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C., Plopper, F., Polindexter, A., Popovic, D., Primus, E., Pu, B.-L., Puazo, M., Quirroz, J., Rachlin, E., Reeves, K., Regier, M.A., Reigh, R., Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, P., Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S.J., Sanders, W., Savery, G., Scherer, S., Scott, G., Shatsman, S., Shen, H., Shetty, J., Shvartabeyan, A., Sisson, I., Sitter, C.D., Smajls, D., Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Sosa, J., Steimle, M., Strong, R., Sutton, A., Svatek, A., Tabor, P., Taylor, C., Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umani, K., Valas, R., Vera, V., Villasana, D., Waldron, L., Walker, B., Wang, J., Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, P., Williams, G., Willson, R., Wlarczyk, R., Wooden, H., Worley, K., Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V., Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Dunn, D., von Niederhausen, A., Weiss, R., Smith, D.R., Holt, R.A., Smith, H.O., Weinstock, G. and Gibbs, R.A.

TITLE
JOURNALREFERENCE
AUTHORSTITLE
JOURNAL

COMMENT

Unpublished
2 (bases 1 to 238039)
Rat Genome Sequencing Consortium.

Direct Submission

Submitted (20-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (<http://www.hgsc.bcm.tmc.edu/projects/rat/>). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: K2X1

Center clone name: CH230-unknown

----- Summary Statistics

Assembly program: Phrap; version 0.990329

Consensus quality: 191709 bases at least Q40

Consensus quality: 196023 bases at least Q30

Consensus quality: 198803 bases at least Q20

Estimated insert size: 196123; sum-of-contigs estimation

Quality coverage: 5x in Q20 bases; sum-of-contigs estimation

* NOTE: Estimated insert size may differ from sequence length (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

* 1 7602: contig of 7602 bp in length
* 7603 7702: gap of unknown length
* 7703 235224: contig of 227522 bp in length
* 235225 235224: gap of unknown length
* 235225 236488: contig of 1164 bp in length
* 236489 236588: gap of unknown length
* 236589 238039: contig of 1451 bp in length.

FEATURES

source

1..238039

/organism="Rattus norvegicus"

/mol_type="genomic DNA"

/db_xref="taxon:10116"

/clone="CH230-unknown"

165878..167444

/note="wgs_contig"

233194..235224

/note="wgs_contig"

ORIGIN

Query Match	41.5%	Score	1156.2	DB 2	Length	238039
Best Local Similarity	99.7%	Pred.	No. 5.1e-267			
Matches	1158	Conservative	0	Mismatches	3	Indels
						Gaps
						0
QY	189	GGTACCCAGGACTGATTCCTCGAAGATGTTCTAGGCTGCAGAAATCTC	CACGCGCACAAGAGT	248		
Db	104379	GGTACCCAGGACTGATTCCTCGAAGATGTTCTAGGCTGCAGAAATCTC	CACGCGCACAAGAGT	104320		
QY	249	TTGGAGTCACAGGATGGGTGCTCGCAAGAGCTAGGACAGAACTGTTGT	CAGCCCTGT	308		
Db	104319	TTGGAGTCACAGGATGGGTGCTCGCAAGAGCTAGGACAGAACTGTTGT	CAGCCCTGT	104260		
QY	309	GCCCGGACCTGTGAGCTGTGAGAAAGACAGAGTCCCACCCAGGCCCTT	CTTAGACCCA	368		
Db	104259	GCCCGGACCTGTGAGCTGTGAGAAAGACAGAGTCCCACCCAGGCCCTT	CTTAGACCCA	104200		
QY	369	CCCCGGGTTTCCAGCATCTTCTTCGAGGACCGGACCCCTGGCTGAAAG	TACAGAAAC	428		
Db	104199	CCCCGGGTTTCCAGCATCTTCTTCGAGGACCGGACCCCTGGCTGAAAG	TACAGAAAC	104140		
QY	429	CCTAGAGTCGAGCCCATGTGGCTGGGCGGCATGTTTCCAGAAATCCTCT	GTGTAAGG	488		
Db	104139	CCTAGAGTCGAGCCCATGTGGCTGGGCGGCATGTTTCCAGAAATCCTCT	GTGTAAGG	104080		
QY	489	ATCCAGACTCTTTCAGGAGCCCAACAGCTCAAGGACAGTTAGCATGTTT	CATGTGTACTG	548		
Db	104079	ATCCAGACTCTTTCAGGAGCCCAACAGCTCAAGGACAGTTAGCATGTTT	CATGTGTACTG	104020		
QY	549	GGAGGACGAGGAGCCCAACAGAGTCTATGAAGATCCACAGGGGCTCGGT	TTCGAGGCCCT	608		
Db	104019	GGAGGACGAGGAGCCCAACAGAGTCTATGAAGATCCACAGGGGCTCGGT	TTCGAGGCCCT	103960		
QY	609	TGGGTTTTATCACCAAAATGTTTCCCAACAGCAATAAACAGCTCCTC	TCAGACAGCGCA	668		
Db	103959	TGGGTTTTATCACCAAAATGTTTCCCAACAGCAATAAACAGCTCCTC	TCAGACAGCGCA	103900		
QY	669	GTGGACCAAGTGGACCAACAGGACAGATCACTCTGTGGGCCAGACTCAT	AGTAACCTCT	728		
Db	103899	GTGGACCAAGTGGACCAACAGGACAGATCACTCTGTGGGCCAGACTCAT	AGTAACCTCT	103840		
QY	729	AACCTCAATCTCCAGCTCCCAAGTCAATGTGGTCACTTGTGTTTCTC	CAGCCACAC	788		
Db	103839	AACCTCAATCTCCAGCTCCCAAGTCAATGTGGTCACTTGTGTTTCTC	CAGCCACAC	103780		


```
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/J"
/db_xref="taxon:10090"
1..5445
join(5446..5489,6209..>6298)
/product="ventricular myosin regulatory light chain"
5446..5489
/number=1
join(5487..5489,6209..>6298)
/codon_start=1
/product="ventricular myosin regulatory light chain"
/protein_id="AAG40240.1"
/db_xref="GI:11761911"
/translation="MAPKKKRIEGSSNVFPEQTIQIFKE"
6209..6298
/number=2

promoter
mRNA
exon
CDS

Query Match
Best Local Similarity 68.8%; Pred. No. 2.9e-94;
Matches 884; Conservative 0; Mismatches 136; Indels 265; Gaps 10;

QY 1150 AAGTCAGCCCTGCTACATAATGAGTTCAAGGTCGCTGGGCTACATGAGACCATGCTT 1209
DB 4426 AGGCTATCTCAGCTACATATCAAGTTCAA-GCCAGCCTGGGCTACATGAGACCCCTGCCT 4484
QY 1210 C-AAAAGAAAAGGAATTGGTAGTATGACATATCTGGTCTCCAGTACTTAGGGACAC 1268
DB 4485 CAAAAGAAAAGGAGCCAGTAGTATGACATATCTCCGCTCTCCAGCACTT-GGGAGAC 4543
QY 1269 AGAGGCCACTCCACACCACTCTCCAGCAGTGGCCTGCTCCCGAGCCTCGTGTATTTC 1328
DB 4544 AGAGGCTACTCCACTGCTCTCCAGCAGCCGGCTGCTCTCCAGCCTCATTTTTTC 4603
QY 1329 ATATCAATGAGATGGGGAACCAACTGCTAAGGTGACCTTTGCAACCCACGGGTGACTGGAG 1388
DB 4604 ATA-----ACATGGGAGCCAACTGCTAAGGTGACCTTTGCTCCCATGGGTGACTGGAG 4657
QY 1389 ACTTGAGAGTGAGGGTATTATCTTCTCCAGTCGGTCAGCAAGTGGTGCCTCCCAAGAA 1448
DB 4658 ACTTGAGAGTGAGTGGTATTATCTTGTCCAGTCTGTGAACAAATGGCAGCTTCAAGGT 4717
QY 1449 GGTTTTGAGTTCAAAGTAGAAGATGGGACAGAGGAGACAGCGAAGAACCCACCTT- 1507
DB 4718 GGTTTTGTGTTCAAAGAGGACATGGACAGGAGAGGCGGAGAGAGCCACCTC 4777
QY 1508 -GGAGCTGACTGTCCCTGTGCGGCTGGGTGGGGACACAAAGCAGAGAACAGAGGCAGAG 1566
DB 4778 AGGAGTAGGCTGTCCCGGTGAAGCTGGGTGGGGACAAAGAGCAGAGAGCAGAGCAGAG 4837
QY 1567 AACAGGTTGGGTGACATTGACAGAGGATGGGTGTCAGAGGCTGCCAAGATGCA 1626
DB 4838 GACAAAGTGGGTGACATTGACAAAGATGGGAATGTCCAGAGGCTGCCAAGATGTG 4897
QY 1627 TAGTGCAAGAGCCCTGAGGTTCCGAGATGCCCTGGATCCGGAATC-AAAGCTCAGGCTCC 1685
DB 4898 CATGTGCAAGAGCCCTGAGGTGCAAGGTTGCTGGATCCAGAGCCAAAGCTCAGGCTC- 4956
QY 1686 TCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1745
DB 4957 ----- 4956
QY 1746 CCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1805
DB 4957 ----- 4956
QY 1806 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1865
DB 4957 -CTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5015
QY 1866 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1925
|||||
```

ORIGIN

```
5016 TCCTCCCC----- 5023
1926 CTCTTCTCTCTCTGCCCCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 1985
5024 ----- 5023
1986 CTCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2045
5024 -----CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCC 5072
2046 TCCTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2105
5073 TCCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 5132
2106 TCCTCCCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2165
5133 TCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5192
2166 CAGAGCATCTGTTCCAGGCCAGGCCAGCCACTGTCTCTTTAACTTTGAAGGGCATTTTTT 2225
5193 CACCCCATCATCCCGAGGCCAGGCCAGCCACTGACTCTTTTAACCTTTGAAGGGCATTTTT 5252
2226 GGTCTCAGCTGTCCACCCAGCCGGGTGTCGACTTTTGAA CGGCTCTTACTTCAGAGAA 2285
5253 GGTCTCAGCTGTCCACCCAGCCGGGTGCGCCCTTTGAGCAGCTCTTACTTCAGAGAA 5312
2286 CGGCATGGGTG-----GGGGGCTTAGGTGGCTCTGCTCACCTACAACTGCCAAA 2339
5313 CGCATGAGTGGGGGTGGGGGCTTAGGTGGCTCTGCTCACCTACAACTGCCAAA 5372
2340 GTGCTCATGGGTATTTTTAAACCCAGGGAAGAGGTATTTATTTCCACAGCAGGGGC 2399
5373 GTGCTCATGGGTATTTTTAAACCCAGGGAAGAGGTATTTATTTCCACAGCAGGGGC 5432
2400 --CGSCCAGCAGGCTCTTGAATTC 2422
5433 AGAGCCAGCAGGCTCTCGAACTC 5457

RESULT 11
AF326769 2939 bp DNA linear ROD 08-MAR-2001
LOCUS Mus musculus myosin light chain 2v gene, promoter region and
partial cds.
AF326769
AF326769.1 GI:13249164
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 2939)
AUTHORS John,M.C. and Powers,P.A.
TITLE Sequence of the mouse myosin light chain 2V promoter
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2939)
AUTHORS John,M.C. and Powers,P.A.
TITLE Direct Submision
JOURNAL Submitted (06-DEC-2000) Physiology, University of Wisconsin,
Madison, 425 Henry Mall, Madison, WI 53706, USA
LOCATION/Qualifiers
FEATURES
source
1..2939
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="129/SV"
/db_xref="taxon:10090"
1..2897
/note="myosin light chain 2v"
2898..>2939
/product="myosin light chain 2v"
2898..>2939
/number=1
2937..>2939

promoter
mRNA
exon
CDS
```


[illegible]

NOTICE: This sequence may not represent the entire insert of this

clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone, fosmid clone or direct clone walk sequence. Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to obtain the consensus sequence; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:

The RPCI-24 BAC Library has been constructed by Pieter de Jong and coworkers (<http://www.chori.org>) from male C57BL/6J mouse spleen and/or brain genomic DNA. The clone and detailed information can be obtained from Pieter de Jong and coworkers at <http://www.chori.org>

NEIGHBORING SEQUENCE INFORMATION:

This sequence is the entire insert of the clone. This clone is overlapped by AC124189.

FEATURES

source	Location/Qualifiers
	1. .161652
	/organism="Mus musculus"
	/mol_type="genomic DNA"
	/db_xref="taxon:10090"
	/chromosome="5"
	/map="5"
	/clone="RP24-67B1"
	/clone_lib="RPCI-24"
repeat_region	116. .195
	/rpt_family="L1"
repeat_region	347. .458
	/rpt_family="L1"
repeat_region	514. .656
	/rpt_family="Alu"
repeat_region	1715. .1997
	/rpt_family="B4"
repeat_region	2293. .2425
	/rpt_family="Alu"
repeat_region	2575. .2632
	/rpt_family="Alu"
repeat_region	2595. .2691
	/rpt_family="B4"
repeat_region	2675. .2746
	/rpt_family="B4"
repeat_region	2776. .2910
	/rpt_family="B4"
repeat_region	3145. .3186
	/rpt_family="B2"
repeat_region	3187. .3329
	/rpt_family="B4"
repeat_region	4125. .4198
	/rpt_family="ID"
repeat_region	4208. .4329
	/rpt_family="B4"
repeat_region	4934. .5102
	/rpt_family="B4"
repeat_region	5225. .5376
	/rpt_family="MER1_type"
repeat_region	5388. .5674
	/rpt_family="MER2_type"
repeat_region	5746. .6144
	/rpt_family="MaLR"
repeat_region	6264. .6555
	/rpt_family="B4"
	/rpt_family="B4"
	/rpt_family="MER2_type"
	6841. .7015
	/rpt_family="B2"
repeat_region	7293. .7499
	/rpt_family="MIR"
repeat_region	7622. .7829
	/rpt_family="MaLR"
repeat_region	8083. .8152
	/rpt_family="Alu"
repeat_region	8087. .8206
	/rpt_family="B4"
repeat_region	8725. .8919
	/rpt_family="B2"
repeat_region	9074. .9113
	/rpt_family="MIR"
repeat_region	9817. .9992
	/rpt_family="B4"
repeat_region	11034. .11228
	/rpt_family="B2"
repeat_region	11823. .11994
	/rpt_family="B4"
repeat_region	12050. .12244
	/rpt_family="B2"
repeat_region	12943. .13127
	/rpt_family="B4"
repeat_region	13684. .13741
	/rpt_family="MER1_type"
repeat_region	13825. .14027
	/rpt_family="B4"
repeat_region	15731. .15834
	/rpt_family="Alu"
repeat_region	15891. .16138
	/rpt_family="B4"
repeat_region	16287. .16499
	/rpt_family="L1"
repeat_region	16774. .16905
	/rpt_family="B4"
repeat_region	19192. .19357
	/rpt_family="B4"
repeat_region	19409. .19497
	/rpt_family="Alu"
repeat_region	20187. .20292
	/rpt_family="MIR"
repeat_region	20790. .20923
	/rpt_family="RMERIA"
repeat_region	24371. .24501
	/rpt_family="Alu"
repeat_region	24802. .24909
	/rpt_family="MIR"
repeat_region	25273. .25622
	/rpt_family="ERVK"
repeat_region	26563. .26594
	/rpt_family="MIR"
repeat_region	26627. .26726
	/rpt_family="MIR"
repeat_region	27965. .28241
	/rpt_family="B4"
repeat_region	28685. .28806
	/rpt_family="MIR"
repeat_region	28846. .28973
	/rpt_family="Alu"
repeat_region	29036. .29226
	/rpt_family="B2"
repeat_region	29668. .29816
	/rpt_family="L1"
repeat_region	30862. .31309
	/rpt_family="ERVK"
repeat_region	32195. .32161
	/rpt_family="MER1_type"
repeat_region	32420. .32626
	/rpt_family="B4"

```

repeat_region 32640..32753 /rpt_family="B2"
repeat_region 32754..32888 /rpt_family="B4"
repeat_region 32889..32980 /rpt_family="L1"
repeat_region 34787..34901 /rpt_family="L2"
repeat_region 35087..35216 /rpt_family="MIR"

Query Match 15.4%; Score 429.6; DB 10; Length 161652;
Best Local Similarity 79.8%; Pred. No. 2e-92;
Matches 597; Conservative 0; Mismatches 139; Indels 14; Gaps 7;

QY 1150 AAGTCAGCCCTGCTACATAATGAGTTCAAGGCTCGCGTGGGTACATGAGACCATGCTT 1209
Db 119287 AGCTATCTCAGCTACATATCAAGTTCAA-GCAGCCTGGGTACATGAGACCATGCTT 119345
QY 1210 C-AAAAAGAAAGAAATGGTATAGTACATATCTGGTCTCTCCAGTACTTAGGGACAC 1268
Db 119346 CAAAAAGAAAGAGGAGCAGTGTAGGACATATCTCCGTCCTCCAGCACTT-GGGAGAC 119404
QY 1269 AGAGCCACTCCACCACTCTCCAGCAGTGGCTGCTCCCGAGCCTCGTTATTTTC 1328
Db 119405 AGAGGCTACTCCACTGCTGCTCCAGCAGCGCGCTGCTCCCTGAGCCTCATTTTTTC 119464
QY 1329 ATATCAATGAGTGGGACCACTGCTAAGGTGACTTGACCCACGCGGTGACTGGAG 1388
Db 119465 ATA-----ACATGGGAGCACTGCTAAGGTGACTTGCTCCATGGGTGACTGGAG 119518
QY 1389 ACTTGAGAGTGGAGGTTTATCATTTTCCAGTCGGTCAGCAAGTGGTCCGCCCAAGAA 1448
Db 119519 ACTTGAGAGTGAAGTGTTATCATTTTGTCCAGTCTGTGAACAAATGGCAGCCTCCAAGT 119578
QY 1449 GGTTTGAGTTCAAAGTAGAAGTGGACAGGAGAGACAGCGAAGACCCACCTT- 1507
Db 119579 GGTTTTGTGTTCAAAGAGGAGACATGGACAGGAGAGCGGAGGAGGAGCCACCCCTC 119638
QY 1508 -GAGCTGACTGTCCTGTGCGCTGGTGGGACACAAAGCAGAGAGGAGGAGGAGAG 1566
Db 119639 AGAGTAGGCTGTCCCGTGAAGCTGGTGGGACAAAGAGCAGAGAGGAGGAGGAGAG 119698
QY 1567 AACAAAGGTGGGTGACATTTTGACAAAGATGGGGGTGTGCGAGAGGCTGCCAAGATGCA 1626
Db 119699 GACAAGGCTGGGTGACATTTTGACAAAGATGGGAATGTGCCAGAGGCTGCCAAGATGTG 119758
QY 1627 TAGGTGCAAGAGGCTGAGGTTGAGGATGCGGTGATCCGGATCAAAGCTCAGGCTCCT 1686
Db 119759 CATGTGCAAGAGGCTGAGGTGCAAGGAGGCTGGATCCAGAGCCAAAGCTCAG--GCT 119816
QY 1687 CCCTCTTCTCTCTCTCTGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTG 1746
Db 119817 CCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 119876
QY 1747 CCCT-CTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1805
Db 119877 TCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 119936
QY 1806 TCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1865
Db 119937 CCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 119996
QY 1866 TCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1995
Db 119997 TCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120026

```

RESULT 13

```

AC111089/c
LOCUS 211580 bp DNA linear HTG 07-OCT-2004
DEFINITION Mus musculus chromosome 5 clone RP23-156F20 map 5, *** SEQUENCING
IN PROGRESS ***, 2 ordered pieces.

```

AC111089

```

AC111089.20 GI:53850722
HTG: HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
Mus musculus (house mouse)

```

ORGANISM

```

Mus musculus
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

REFERENCE

```

1 (bases 1 to 211580)
Birren,B., Nusbaum,C. and Lander,E.

```

AUTHORS

```

Mus musculus chromosome 5, clone RP23-156F20

```

JOURNAL

```

Unpublished

```

REFERENCE

```

2 (bases 1 to 211580)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,

```

AUTHORS

```

Anderson,S., Barna,N., Bastien,V., Boguslavskiy,L., Boukhgalter,B.,
Brown,A., Camarata,J., Campotano,A., Chang,J., Chazaro,B.,
Choepe,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,
Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Horcon,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R.,
Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C.,
Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M.,
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,
Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S.,
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Struss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

```

Direct Submission

```

Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

JOURNAL

```

3 (bases 1 to 211580)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,

```

REFERENCE

```

Anderson,M., Anderson,S., Arachchi,H.M., Barna,N., Bastien,V.,
Bloom,T., Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J.,
Choepe,Y., Collamore,A., Cook,A., Cooke,P., Corum,B.,
DeArellano,K., Diaz,J.S., Dodge,S., Dooley,K., Dorris,L.,
Erickson,J., Faro,S., Ferreira,P., Fitzgerald,M., Gage,D.,
Galagan,J., Gardyna,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagopian,D., Hagos,B., Hall,J., Horton,L., Hulme,W., Iliev,I.,
Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T.,
Levine,R., Lindblad-Toh,K., Liu,G., Liu,X., Lui,A., Mabbitt,R.,
Mlenga,V., Macdonald,P., Major,J., Manning,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nguyen,T., Nicol,R., Norbu,C.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Phunkhang,P., Pierre,N., Rachupka,A., Ramasamy,U., Raymond,C.,
Retta,R., Rise,C., Rogov,P., Roman,J., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Stubbs,M., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Vassiliev,H., Venkataraman,V.S., Viel,R.,
Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L.,
Zimmer,A. and Zody,M.

```

JOURNAL

```

Submitted (07-OCT-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

```

COMMENT

```

On Oct 7, 2004 this sequence version replaced g1:52627295.
All repeats were identified using RepeatMasker:

```

```

Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

```

```

----- Genome Center
Center: Whitehead Institute/MIT Center for Genome Research

```

```

Center code: WIBR

```

```

Web site: http://www-seq.wi.mit.edu

```

```

Contact: sequence_submissions@broad.mit.edu

```

```

----- Project Information

```

```

Center project name: L22447

```


REFERENCE	2 (bases 1 to 4672)	ORIGIN
AUTHORS	Maggi, A.	
TITLE	Direct Submission	
JOURNAL	Submitted (09-MAY-2000) Maggi A., Institute of Pharmacological Sciences, University of Milan, Via Balzaretti 9, Milan I-20133, ITALY	
FEATURES	Location/Qualifiers	
source	1..4672	QY 2417 GAATTCGACCCCTTCGAGCTTGGCATTCGGGTACTGTGTGTTAAATGGAAGACGCCAAA 2476
	/organism="Cloning vector pMAR"	
	/mol_type="other DNA"	
	/db_xref="taxon:125631"	Db 254 GATCTCGAGGCATGCAAGCTTGGCATTCGGGTACTGTGTGTTAAATGGAAGACGCCAAA 313
	/focus	
	/note="artificial sequence"	QY 2477 ACATAAAGAAAGGCCCGCGGCATTCTATCTCTAGAGATGGAAACCGCTCGAGAGCCAA 2536
source	1..39	
	/organism="synthetic construct"	Db 314 ACATAAAGAAAGGCCCGCGGCATTCTATCTCTAGAGATGGAAACCGCTCGAGAGCCAA 373
	/mol_type="other DNA"	
	/db_xref="taxon:32630"	QY 2537 CTGCATAAGGCTAGAAAGAGATACGCCCTTGGTTCCTGGAAACAATTCGTTTTACAGATGCAC 2596
	/note="estrogen receptor responsive elements BRE"	
source	40..79	Db 374 TGCATAAGGCTATGAAGAGATACGCCCTTGGTTCCTGGAAACAATTCGTTTTACAGATGCAC 433
	/organism="synthetic construct"	
	/mol_type="other DNA"	QY 2597 ATATCGAGGTGAACATCAGTTCCGGGAATCTTCGAAATGTCGTTTCGTTGGCAGAA 2656
	/db_xref="taxon:32630"	
	/note="synthetic polylinker"	Db 434 ATATCGAGGTGAACATCAGTACGCCGAATATCTTCGAAATGTCG - TTCGGTTGGCAGAA 492
source	80..247	QY 2657 GCTATGAAACGATATGGGCTGAATAACAATCACAGAATCGTCGTATGCAGTGAATACTCT 2716
	/organism="Herpes simplex virus unknown type"	
	/mol_type="other DNA"	Db 493 GCTATGAAACGATATGGGCTGAATAACAATCACAGAATCGTCGTATGCAGTGAATACTCT 552
	/db_xref="taxon:126283"	
	248..297	QY 2717 CTTTCAATTCCTTTATGCCGGTGTGGGCCCGCTATTTATTCGCGAGTTGCAGTTCCGCC 2775
source		
	/organism="synthetic construct"	Db 553 C-TTCAATTCCTTTATGCCGGTGTGGGCCCGCTATTTAT - CGGAGTTGCAGTTCCGCC 609

Search completed: August 1, 2005, 15:56:54
Job time : 11863 secs

This Page Blank (nsptot)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2005, 11:32:52 ; Search time 1436 Seconds
(without alignments)
11476.713 Million cell updates/sec

Title: US-09-068-751-1

Perfect score: 2784

Sequence: 1 gaagtgaatctgaataatt.....cagttgccgcgcgcgaaca 2784

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04.*

1: Geneseqn1980s.*

2: Geneseqn1990s.*

3: Geneseqn2000s.*

4: Geneseqn2001as.*

5: Geneseqn2001bs.*

6: Geneseqn2002as.*

7: Geneseqn2002bs.*

8: Geneseqn2003as.*

9: Geneseqn2003bs.*

10: Geneseqn2003cs.*

11: Geneseqn2003ds.*

12: Geneseqn2004as.*

13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2760	99.1	2784	2 AAT65102	Aat65102 Luciferas
C 2	393.6	14.1	110000	11 ACN43984_1	Continuation (2 of
C 3	393	14.1	110000	11 ACN43984_2	Continuation (3 of
4	284.6	10.2	4672	6 ABS55458	Chicken 1
5	283.8	10.2	5819	2 AAX08774	Aax08774 pGL2(apo
6	283.8	10.2	5819	2 AAX08775	Aax08775 pGL2(apo
7	283.8	10.2	5819	3 AAZ38351	Aaz38351 Plasmid p
8	283.8	10.2	6092	2 AAT85710	Aat85710 Human apo
9	283.8	10.2	6092	3 AAX08773	Aax08773 pGL2(apo
10	283.8	10.2	6092	3 AAZ38350	Aaz38350 Plasmid p
11	282.2	10.1	5819	2 AAT85711	Aat85711 Human apo
12	280.2	10.1	11616	2 AAQ38105	Aaq38105 vector pH
13	280.2	10.1	11616	2 AAQ67390	Aaq67390 Construct
14	279.8	10.1	5789	2 AAX08779	Aax08779 pGL2-Prom
15	279.8	10.1	5791	2 AAX08778	Aax08778 pGL2 (GST-
16	279.8	10.1	5793	2 AAX08777	Aax08777 pGL2(apo
17	279.8	10.1	5793	2 AAX08776	Aax08776 pGL2(apo
18	279.8	10.1	5818	3 AAZ38352	Aaz38352 Plasmid p
19	279.8	10.1	5849	12 ADO43719	Ado43719 Nucleotid
20	279.8	10.1	5938	2 AAT85713	Aat85713 Human apo

21	279.8	10.1	5938	3 AAZ38353	Aaz38353 Plasmid p
22	279.8	10.1	8937	4 AAH22812	Aah22812 DNA seque
23	279.8	10.1	8937	4 AAH22811	Aah22811 DNA seque
24	279.8	10.1	8937	10 ABZ58699	Abz58699 DNA seque
25	279.8	10.1	8937	10 ABZ58700	Abz58700 DNA seque
26	279.8	10.1	13940	5 AAF25915	Aaf25915 Human tum
27	278.4	10.0	1811	2 AAQ58732	Aaq58732 Encodes f
28	278.4	10.0	1838	6 ABK85491	Abk85491 cDNA enco
29	278.4	10.0	5620	2 AAT48630	Aat48630 Ad5-ITR,
30	278.4	10.0	5620	4 AAF30233	Aaf30233 Minimal a
31	278.2	10.0	5818	2 AAT85712	Aat85712 Human apo
32	275.8	9.9	6170	2 AAQ98929	Aaq98929 Vector p1
33	275.8	9.9	6229	2 AAQ98931	Aaq98931 Vector p3
34	275.8	9.9	6254	2 AAQ98918	Aaq98918 TGF-beta
35	275.8	9.9	6254	2 AAQ98916	Aaq98916 TGF-beta
36	275.8	9.9	6464	2 AAQ98933	Aaq98933 Vector HB
37	275.8	9.9	10533	2 AAQ98930	Aaq98930 Vector p3
38	275.8	9.9	10549	2 AAQ98911	Aaq98911 TGF-beta
39	275.8	9.9	10558	2 AAQ98912	Aaq98912 TGF-beta
40	275.8	9.9	10558	2 AAQ98914	Aaq98914 TGF-beta
41	275.8	9.9	10569	2 AAQ98913	Aaq98913 TGF-beta
42	275.8	9.9	10697	2 AAQ98910	Aaq98910 TGF-beta
43	275.8	9.9	10768	2 AAQ98932	Aaq98932 Vector HB
44	275.8	9.9	11293	2 AAQ98909	Aaq98909 TGF-beta
45	274.8	9.9	7223	2 AAT62602	Aat62602 Luciferas

ALIGNMENTS

RESULT 1

AAT65102

ID AAT65102 standard; DNA; 2784 BP.

XX

AC AAT65102;

XX

DT 25-MAR-2003 (revised)

DT 24-FEB-1998 (first entry)

XX

DE Luciferase gene under control of rat MLC-2 gene promoter.

XX

KW Luciferase; upstream regulatory region; MLC-2 promoter; gene therapy;

KW rat cardiac myosin light chain gene; heart disease; cardiac muscle; ss.

XX

OS Synthetic.

OS Rattus sp.

OS unidentified adenovirus.

OS Chimeric.

XX

Key Location/Qualifiers

FT misc_feature 1..156

FT promoter /tag= a

FT /note= "encodes Adenovirus Ad5 psi encapsidation"

FT /tag= b

FT /note= "rat MLC-2v gene promoter"

FT misc_signal 682..724

FT /tag= c

FT /note= "CSS-specific sequence"

FT misc_signal 2207..2219

FT /tag= d

FT /note= "HF-3 element"

FT misc_signal 2229..2241

FT /tag= e

FT /note= "MLE1 element"

FT misc_signal 2271..2289

FT /tag= f

FT /note= "HF-2 element"

FT misc_signal 2328..2333

FT /tag= g

FT /note= "E-box element"

FT misc_signal 2340..2348

FT /tag= h

```
FT misc_signal /note= "HF-1a element"  
FT 2349. .2361  
FT /tag= i  
FT /note= "HF-1b element"  
FT prim_transcript 2406. .2784  
FT /tag= j  
FT /note= "partial primary transcript"  
FT CDS 2461. .2784  
FT /tag= k  
FT /product= "luciferase"  
FT /note= "partial sequence"  
XX  
PN W09717937-A2.  
XX  
XX 22-MAY-1997.  
XX  
XX 14-NOV-1996; 96WO-DE002181.  
XX  
XX 17-NOV-1995; 95DE-01042838.  
PR 01-OCT-1996; 96DE-01040630.  
XX  
XX (FRAN/) FRANZ W.  
XX  
XX Franz W, Rothmann T, Katus HA;  
PI WPI; 1997-289035/26.  
XX  
XX Nucleic acid construct for gene therapy of cardiac diseases - contains  
PT promoter from myosin light chain gene to provide heart-specific  
PT expression of therapeutic protein, antisense sequence or ribozyme.  
XX  
XX Disclosure; Fig 10A-C; 44pp; German.  
XX  
XX This sequence contains part of the luciferase coding sequence under the  
CC control of the upstream regulatory region from the rat cardiac myosin  
CC light chain gene (MLC-2) which when inserted in viral vector  
CC pAD.RSVbetagal (in place of the Rous sarcoma virus/beta-galactosidase  
CC gene cassette) creates pAD-mluc. Nucleic acid constructs containing a  
CC regulatory sequence from the 5'-end of the MLC-2 gene functionally  
CC attached to a sequence encoding a therapeutic product, antisense molecule  
CC or ribozyme, are useful in gene therapy. Preferred therapeutic sequences  
CC encode dystrophin, a beta-adrenergic receptor, nitric oxide synthase, and  
CC may include one or more non-coding regions and/or a polyA signal. The  
CC constructs are used to treat heart disease, specifically cardiac  
CC insufficiency, dilative or hypertrophic cardiomyopathy, dystrophin-  
CC related disease, vascular disease, hypertension, arteriosclerosis,  
CC stenosis and/or restenosis. In particular the constructs provide high  
CC transfection rates, stable gene expression and are specific for cardiac  
CC muscle. (Updated on 25-MAR-2003 to correct PR field.)  
XX  
XX Sequence 2784 BP; 571 A; 867 C; 635 G; 711 T; 0 U; 0 Other;  
SQ  
  
Query Match 99.1%; Score 2760; DB 2; Length 2784;  
Best Local Similarity 99.5%; Pred. No. 0;  
Matches 2769; Conservative 0; Mismatches 15; Indels 0; Gaps 0;  
  
QY 1 GAAGTGAATCTGAATAATTTTGTTGTTACTCATAGCGGTAAATATTGTTCTAGGCGCGG 60  
DB 1 GAAGTGAATCTGAATAATTTTGTTGTTACTCATAGCGGTAAATATTGTTCTAGGCGCGG 60  
  
QY 61 GGACTTTGACCGTTTACGTGGAGACTCGCCCGAGGTGTTTTTCTCAGGTGTTTTTCGCGGTT 120  
DB 61 GGACTTTGACCGTTTACGTGGAGACTCGCCCGAGGTGTTTTTCTCAGGTGTTTTTCGCGGTT 120  
  
QY 121 CCGGGTCAAAAGTTGGCGTTTATTATATAGTCAGGGGATCCGGAATCTTTGAAGACGA 180  
DB 121 CCGGGTCAAAAGTTGGCGTTTATTATATAGTCAGGGGATCCGGAATCTTTGAAGACGA 180  
  
QY 181 AAGGGCCCGGTACCCAGGACTGATCTCGAAAGTTCTAGGTCGAGAATCTCACACGC 240  
DB 181 AAGGGCCCGGTACCCAGGACTGATCTCGAAAGTTCTAGGTCGAGAATCTCACACGC 240  
  
QY 241 ACAAGAGTTTGGAGTCACAGGATGGGTGCTCGCCAAAGAGCCTAGGGAAGAAGCTTGTCA 300
```

```
DB 241 ACAAGAGTTTGGAGTCACAGGATGGGTGTCCGCCAAGAGCCTTAGGGACAGACGTTGTCA 300  
QY 301 GCCCTGTGCCCGGACCCCTGTGGACTGTGAGAAGACAGAGTCCACCCAGGCTTCT 360  
DB 301 GCCCTGTGCCCGGACCCCTGTGGACTGTGAGAAGACAGAGTCCACCCAGGCTTCT 360  
QY 361 TAGACCCACCCCGGGTTTTTCCAGCATCTTCTCGAGGACCGGACCCCTGGCTGAAAGT 420  
DB 361 TAGACCCACCCCGGGTTTTTCCAGCATCTTCTCGAGGACCGGACCCCTGGCTGAAAGT 420  
QY 421 ACAGAAAACCTTAGAGTCTGACGCCCATGTGTGGCGCCCATGTTTCCAGAAATCCTCTG 480  
DB 421 ACAGAAAACCTTAGAGTCTGACGCCCATGTGTGGCGCCCATGTTTCCAGAAATCCTCTG 480  
QY 481 GTCTAAGGATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGGACAGTAGCATGTTTCA 540  
DB 481 GTCTAAGGATCCAGACCTCTTACGGAGCCCAACAGCTCAAGGGACAGTAGCATGTTTCA 540  
QY 541 GTGTACTGGGAGGACGAGGACCAACAGAGGTCAATGAAGATCCACAGGGGCTCCGGTTCC 600  
DB 541 GTGTACTGGGAGGACGAGGACCAACAGAGGTCAATGAAGATCCACAGGGGCTCCGGTTCC 600  
QY 601 GAGGCCCTTGGGTTTTTATCAACAAATGTTTCCACCAGCAACATAAAACAGCTCCTCAG 660  
DB 601 GAGGCCCTTGGGTTTTTATCAACAAATGTTTCCACCAGCAACATAAAACAGCTCCTCAG 660  
QY 661 ACAGCGCAGTGGACAGTGGACCAAGGACAGATCACTCTGTGGGCCAGACATCATAG 720  
DB 661 ACAGCGCAGTGGACAGTGGACCAAGGACAGATCACTCTGTGGGCCAGACATCATAG 720  
QY 721 TAACTCTAACTCAATCTCCAGCCTCCACAGTCATTTGTCGGTCACTTGTCTCAGC 780  
DB 721 TAACTCTAACTCAATCTCCAGCCTCCACAGTCATTTGTCGGTCACTTGTCTCAGC 780  
QY 781 CACCACTTGGCAAGTCACTGTGTCCTCAACAAATCTTCAGAAAGCAGGCGGATGGGG 840  
DB 781 CACCACTTGGCAAGTCACTGTGTCCTCAACAAATCTTCAGAAAGCAGGCGGATGGGG 840  
QY 841 TTTTGTGTTAACTGATGGGTGTTTTGTTGTTGTTTCTTAACTGTCACGTAGCCAG 900  
DB 841 TTTTGTGTTAACTGATGGGTGTTTTGTTGTTGTTTCTTAACTGTCACGTAGCCAG 900  
QY 901 GCTAGCCTTGAACCTCACTATGTAGGCAAGCATGACCATGAATCTTGATCTCTCTC 960  
DB 901 GCTAGCCTTGAACCTCACTATGTAGGCAAGCATGACCATGAATCTTGATCTCTCTC 960  
QY 961 AGTGTCTGGGATACAGGTGTGTCTCTCCCTACCTTCTAATAGCAATATGTGGCA 1020  
DB 961 AGTGTCTGGGATACAGGTGTGTCTCTCCCTACCTTCTAATAGCAATATGTGGCA 1020  
QY 1021 CATGTTTGTGCCCCACAGGTTGAGACCATCTTGACCTGAGGAAGAAATAGCTAACTCA 1080  
DB 1021 CATGTTTGTGCCCCACAGGTTGAGACCATCTTGACCTGAGGAAGAAATAGCTAACTCA 1080  
QY 1081 CCTCTGAAGGTGCGCTGGATCTCGTCTTTGCTTTCCAGCACTCAGAGTGGGGGGTTC 1140  
DB 1081 CCTCTGAAGGTGCGCTGGATCTCGTCTTTGCTTTCCAGCACTCAGAGTGGGGGGTTC 1140  
QY 1141 AGAAGTGCAAGTACAGCCCTGCTACATAAATAGAGTTCAAGGTCGCTGGGCTACATGAG 1200  
DB 1141 AGAAGTGCAAGTACAGCCCTGCTACATAAATAGAGTTCAAGGTCGCTGGGCTACATGAG 1200  
QY 1201 ACCATGCTCAAAAAGAAAGAAATTTGGTATAGTCACATCTCTGGTCTCTCCAGTACTT 1260  
DB 1201 ACCATGCTCAAAAAGAAAGAAATTTGGTATAGTCACATCTCTGGTCTCTCCAGTACTT 1260  
QY 1261 AGGGACACAGAGGCCACTCCACCATCTCCAGCAGCTGGGCTCCCTCCCGAGCCTCG 1320  
DB 1261 AGGGACACAGAGGCCACTCCACCATCTCCAGCAGCTGGGCTCCCTCCCGAGCCTCG 1320  
QY 1321 TTTATTTCAATATCAATAGATGGGGAACCAACTGTAAAGGTGACCTTGCAACCCAGGGGT 1380  
DB 1321 TTTATTTCAATATCAATAGATGGGGAACCAACTGTAAAGGTGACCTTGCAACCCAGGGGT 1380
```


QY	1627	TAGGTGCAAGGCGCTGAGGTTGCGAGATGCTGATGATCCGGAATCAAAGCTCAGGCTCCT	1686
Db	100241	CATGTGCAAGGCGCTGAGGTTGCGAGATGCTGATGATCCGGAATCAAAGCTCAG--GCT	100184
QY	1687	CCCTCTTCTCCTCCTCCTCCTGCGCCCTCCTCCTCCTCCTGCGCCCTCCTCCTCCTGCGC	1746
Db	100183	CCCT	100124
QY	1747	CCCTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	1806
Db	100123	TCCTC-----CCCTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	100069
QY	1807	CCCT	1866
Db	100068	CCTCCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	100009
QY	1867	CCTCCCTCCTCATCTACCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	1926
Db	100008	NN	99949
QY	1927	CTCTCCTCCTCTGCGCCCTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	1986
Db	99948	NN	99889
QY	1987	TCCTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	2046
Db	99888	NN	99829
QY	2047	CCTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	2106
Db	99828	NN	99769
QY	2107	CCTCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	2166
Db	99768	NN	99709
QY	2167	AGAGCATCGTCCAGCGCCAGCCAGCCACTGTCTCTTAACTCTGAAGCATTTTGT	2226
Db	99708	NN	99649
QY	2227	GGTCTCAGTGTCCACCGCGGGGTGCGAGATTGACGGCTCTTACTTCAGAGAAC	2286
Db	99648	NN	99589
QY	2287	GGCATGGGTGGGGGGCTTAGTGCCCTCTGCCTCACCTACAACTGCCAAAGTGTCA	2346
Db	99588	NN	99529
QY	2347	TGGGGTTATTTTAAACCCAGGGAAGAGGTATTTATTTGTCCACAGCGGCGCGCCAG	2406
Db	99528	NN	99469
QY	2407	CAGGCTCCTGAATCGACCCCTTGAGCTTGGCATTCGGTACTCGGTAAATGGAA	2466
Db	99468	NN	99409
QY	2467	GAGGCCA 2473	
Db	99408	CACCCCA 99402	

RESULT 3
ACN43984_2/c
Continuation (3 of 5) of ACN43984 from base 200001 (Mouse genomic sequence MCG12182.)
WP Sequence split into 5 fragments LOCUS ACN43984 Accession ACN43984
Fragment Name Begin End
WP ACN43984_0 1 110000
WP ACN43984_1 100001 210000
WP ACN43984_2 200001 310000
WP ACN43984_3 300001 410000
WP ACN43984_4 400001 493631

Query Match 14.1%; Score 393; DB 11; Length 110000;

		Best Local Similarity 81.0%; Pred. No. 2.9e-68; Matches 547; Conservative 0; Mismatches 110; Indels 18; Gaps 7;	
QY	1150	AAGTCAGCCCTGCTACATAATAGTTCAGGCTCGCTGGGCTACATGAGACCATGCT	1209
Db	713	AGGCTATCCTCAGCTACATATCAAGTTCAA--GCCAGCTGGGCTACATGAGACCTGCT	655
QY	1210	C-AAAAAGAAAGGAATTTGGTATAGTGACATACTCTGGTCTCTCCAGTACTTAGGGAAC	1268
Db	654	CAAAAAGAAAGGAGCCAGTGTAGGACATACTCCCTCTCTCCAGCACIT--GGAGAC	596
QY	1269	AGAGGCCATCCACCAACATCTCCAGCAGCTGGCTGCTCCCGAGCTGTTTATTTTC	1328
Db	595	AGAGGCTACTCCACTGCTGTCTCCAGCAGCGGCTGCTCCTGAGGCTCATTTTTTTC	536
QY	1329	ATATCAATAGATGGGACCCAACTGTAAAGTGACCTTGACCCACGGGTGACTGGAG	1388
Db	535	ATA-----ACATGGGGACCCAACTGTAAAGTGACCTTGCTCCCATGGGTGACTGGAG	482
QY	1389	ACTTGAGAGTGGAGGTTTATCATTTCTCCAGTCCGTGAGCAAGTGGTGGCGCCAGAA	1448
Db	481	ACTTGAGAGTGAAGTGGTTATCATTTGTCCAGTCTGTGAAACAATGGCAGCTCCAAGT	422
QY	1449	GTTTGTAGTTCAAAGTAAAGATGGGACAGGAGAGACCCAGCGAGAGAACCCACCT-	1507
Db	421	GTTTGTGTTCAAAGGAGGACATGGGACAGGAGAGGCGGAGAGAGAGAGAGAGAGAG	362
QY	1508	-GGAGCTGACTGCTCCTGCTGCGGCTGGGTGGGACACAAAGCAGAGAGAGAGAGAG	1566
Db	361	AGGAGTAGGCTGTCCCTGGAAGCTGGGTGGGGAACAAAGCAGAGAGAGAGAGAGAG	302
QY	1567	AACAGGCTGGGTGACATTTGAGCAAGGATGGGGGTGTGCCAGAGGCTGCCCAAGATGA	1626
Db	301	GACAGCGTGGGTGACATTTGAGCAAGATGGGAATGTGCCAGAGGCTGCCCAAGATGTG	242
QY	1627	TAGTGAAGAGCCCTGAGGTTGAGGATGCTGTGATCCGGAATCAAAGCTCAGGCTCCT	1686
Db	241	CATGTGCAAGGCCCTGAGGTGCAAGGGTGTGATCCAGAGCCAAAAGGTCTAG--GCT	184
QY	1687	CCCTCTTCTCCTCCTCCTGCGCCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	1746
Db	183	CCCT	124
QY	1747	CCCTCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	1806
Db	123	TCCTC-----CCCTTCTCCTTCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCTCCT	69
QY	1807	CCCTCCTCCTCCTCCT 1821	
Db	68	CCTCCCTTCTCCTCCT 54	

RESULT 4
ABS55458
ID ABS55458 standard; DNA; 4672 BP.
XX
AC ABS55458;
XX
DT 31-DEC-2002 (first entry)
XX
DE Chicken lysozyme matrix attachment region (MAR) element #5.
XX
KW Chicken; eukaryotic cell transfection; nucleic acid vector;
KW chromatin element; matrix attachment region; MAR element;
KW stable eukaryotic cell line; recombinant protein;
KW transient transfection efficiency; lysozyme; ds.
XX
OS Gallus gallus.
XX
FN WO200274969-A2.
XX
PD 26-SEP-2002.
XX

```

PF 28-JAN-2002; 2002WO-IB002137.
XX
PR 26-JAN-2001; 2001US-0264355P.
PR 04-APR-2001; 2001US-0281391P.
XX
XX (UYLA-) UNIV LAUSANNE.
XX
XX Mermoud N, Zahn-Zabal M, Imhof M, Chatellard P, Girod P;
XX WPI; 2002-750557/81.
XX
XX Transforming a eukaryotic cell by co-transfecting or introducing 2
XX PT vectors into a cell, useful for generating cell lines that express one or
XX PT more genes encoding for recombinant proteins.
XX
XX Claim 5; Page 16-18; 52pp; English.
XX
XX The present invention relates to compositions and methods of transfecting
XX CC eukaryotic cells with nucleic acid vectors, and the use of chromatin
XX CC elements such as matrix attachment region (MAR) elements for
XX CC transfection. The methods and compositions of the invention are useful
XX CC for generating stable eukaryotic cell lines that express one or more
XX CC genes encoding for recombinant proteins. The MAR elements are useful in
XX CC increasing stable and transient transfection efficiency. The present
XX CC sequence represents a chicken lysozyme MAR element
XX
XX Sequence 4672 BP; 1262 A; 916 C; 1056 G; 1438 T; 0 U; 0 Other;
SQ
Query Match 10.2%; Score 284.6; DB 6; Length 4672;
Best Local Similarity 92.5%; Pred. No. 6.9e-47;
Matches 332; Conservative 0; Mismatches 24; Indels 3; Gaps 3;
QY 2417 GAATTCACCCCTCGAGCTTGGCATTCCGGTACTGTGGTAAATGGAAGACGCCAAAA 2476
DB 254 GATCTCGAGCATGCAAGCTTGGCATTCCGGTACTGTGGTAAATGGAAGACGCCAAAA 313
QY 2477 ACATAAAGAAAGCCCGGCCCATCTCTATCCCTAGAGATGGAACCGCTGGAGCCAA 2536
DB 314 ACATAAAGAAAGCCCGGCCCATCTCTATCCCTAGAGATGGAACCGCTGGAGCAAC 373
QY 2537 CTGCATAAGGCTAGAGAGATACGCCCTGGTCTCTGGAACAATTGCTTTACAGATGCAC 2596
DB 374 TGCAATAGGCTATGAAGAGATACGCCCTGGTCTCTGGAACAATTGCTTTACAGATGCAC 433
QY 2597 ATATCGAGGTGAACATCACGTTCCGCGAATACCTTCGAAATGTCCGTTTCGGTTGGCAGAA 2656
DB 434 ATATCGAGGTGAACATCACGTTCCGCGAATACCTTCGAAATGTCCG-TTCGTTGGCAGAA 492
QY 2657 GCTATGAACGATATGGGCTGAATACAAATCACAGATGCTGCTATGCAATGCAAACTCT 2716
DB 493 GCTATGAACGATATGGGCTGAATACAAATCACAGATGCTGCTATGCAATGCAAACTCT 552
QY 2717 CTTTCAATTCTTATCCGGTGTGGCCGGTTATTATTCGCGAGTTGCGAGTTGCCGCC 2775
DB 553 C-TTCAATTCTTATCCGGTGTGGCCGGTTATTAT-CGAGTTGCGAGTTGCCGCC 609

RESULT 5
AX08774
ID AAX08774 standard; DNA; 5819 BP.
XX
XX AAX08774;
XX
XX 27-SEP-1999 (first entry)
XX
XX pGL2 (apo AI-250) luc plasmid construct comprising ARE/luciferase.
XX
XX Antioxidant responsive element; ARE; low density lipoprotein; LDL;
XX KW high density lipoprotein; HDL; apolipoprotein; apo AI; atherosclerosis;
XX KW heart disease; transcription; ss.
XX
XX Synthetic.
XX

```

```

PN CA2238662-A.
XX
XX 23-NOV-1998.
XX
XX 22-MAY-1998; 98CA-02238662.
XX
XX 23-MAY-1997; 97US-00862431.
XX (TOOH ) UNIV QUEENS KINGSTON.
XX Tam S;
XX WPI; 1999-229918/20.
XX
XX New Antioxidant Response Element (ARE), useful for identifying drugs and
XX PT transcription factors for increasing transcription of mRNA, useful for
XX PT treatment of atherosclerosis.
XX
XX Example; Page 55-59; 115pp; English.
XX
XX DNA constructs comprising antioxidant responsive elements (AREs) are
XX CC useful for screening for compounds and transcription factors that bind to
XX CC the ARE and increase transcription levels of a mRNA regulated by an ARE.
XX CC AREs may also be useful as a reagent for purification of a compound
XX CC (preferably a transcription factor) with which it interacts. High Density
XX CC lipoprotein (HDL) has antioxidant activity and protects against oxidized
XX CC low-density lipoprotein (LDL) which has a role in the etiology of
XX CC atherosclerosis. Apolipoprotein (apo) AI is a major component of HDL, and
XX CC is believed to promote the process of reverse cholesterol transport. The
XX CC transcription of apo AI is effected by cis- and trans-acting factors (i.e
XX CC an ARE) UV cross-linking studies using an apoAI-ARE probe isolated two
XX CC polypeptides of 100 and 115 kDa. These compounds are useful for treatment
XX CC of a human or animal with atherosclerosis. ARE's can also be used in DNA
XX CC constructs when operably linked to heterologous protein coding sequences
XX CC to effect the transcription of those heterologous sequences. A Genelight
XX CC vector (pGL2-B) comprising a 491 bp fragment of the apo AI from -491 to
XX CC +1 inserted into the XhoI site of the vector and upstream of the
XX CC luciferase gene was designated pGL2 (apoAI-491) luc. Plasmid pGL2 (apoAI-
XX CC 250) luc was constructed by releasing a DNA fragment (-491 to -251 of the
XX CC apo AI promoter) from pGL2 (apoAI-491) luc. A new plasmid pGL2 (apoAI-250
XX CC mutant ARE) was prepared by PCR using four primers (AAX08785-88). Two
XX CC of DNA, the remaining two primers (AAX08787, AAX08788) were complementary
XX CC to the coding strand of DNA. The plasmid pGL2 (apoAI250) luc was used as a
XX CC template and the DNA fragment was purified and then cloned into the SmaI
XX CC and HindIII sites of pGL2-B to generate pGL2 (apoAI-250 mutant ARE). A
XX CC series of pGL2-Promoter vectors comprising apoAI-ARE, GST-ARE and mutated
XX CC ARE were also constructed. See AAX08773-X08779
XX
XX Sequence 5819 BP; 1569 A; 1349 C; 1317 G; 1584 T; 0 U; 0 Other;
SQ
Query Match 10.2%; Score 283.8; DB 2; Length 5819;
Best Local Similarity 90.6%; Pred. No. 1e-46;
Matches 336; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2405 AGCAGCTCTTGAATTCGACCCCTTCGAGCTTGGCATTCCGCTACTGTGGTAAATGG 2464
DB 242 AGCTGCTCTTAGTCGAGATCTAAGTAAGCTTGGCATTCCGCTACTGTGGTAAATGG 301
QY 2465 AAGACGCCAAAACATAAAGAAAGCCCGCGCCATCTCTCTCTAGAGGATGGAACCG 2524
DB 302 AAGACGCCAAAACATAAAGAAAGCCCGCGCCATCTCTCTCTAGAGGATGGAACCG 361
QY 2525 CTGGAGAGCAACTGCATAAGGCTAGAAGAGATACGCCCTGGTTCCTGGAACAATTGCTT 2584
DB 362 CTGGAGAGCAACTGCATAAGGCTAGAAGAGATACGCCCTGGTTCCTGGAACAATTGCTT 421
QY 2585 TTACAGATGCACATATCGAGGTGAACATCACGTTCCCGGAATACTTCGAAATTCGGTTT 2644
DB 422 TTACAGATGCACATATCGAGGTGAACATCACGTTACGCGGAATACTTCGAAATTCGG-TT 480
QY 2645 CGTTGGCAGAGCTATGAACGATATGGCTGAATACAAATCACAGATCGTCGATGCG 2704

```


XX
CC This sequence represents plasmid pGL2 (apoAI-250) luc. This plasmid
CC comprises a 250 bp portion of the 5' flanking region of the human
CC apolipoprotein AI (apo AI) gene inserted upstream of the luciferase
CC reporter gene in the pGL2-Basic vector. The pGL2-Basic vector also
CC comprises a functional polyadenylation sequence downstream of the
CC reporter gene. This plasmid was derived from plasmid pGL2 (apoAI-491) luc
CC (AAZ38351) and used as a prototype for the construction of pGL2 (apoAI-
CC DRE)/TK/luc (AAZ38352) and pGL2 (4xapoAI-DRE)/TK/luc (AAZ38353). pGL2
CC (apoAI-DRE)/TK/luc contains a fragment of the apoAI 5' flanking sequence
CC (bases -79 to -44 from the transcriptional start site) which contains two
CC drug responsive elements (DRES, AAZ38354-Z38356) and an HSV thymidine
CC kinase (TK) promoter. pGL2 (4xapoAI-DRE)/TK/luc is identical, except that
CC it contains 4 copies of the apoAI flanking sequence fragment. Plasmids
CC pGL2 (apoAI-DRE)/TK/luc and pGL2 (4xapoAI-DRE)/TK/luc are used in a novel
CC method for screening for a drug that increases expression of a gene for
CC apolipoprotein (apo AI), which is associated with coronary artery
CC disease. Such plasmids may be transformed into mammalian cells, which are
CC treated with a candidate compound, lysed and assayed for reporter gene
CC activity relative to extracts from untreated cells. The method is useful
CC for screening and identifying drugs that increase apo AI gene expression
CC the identified drugs would be useful for treating coronary artery
CC disease. The method is simple, rapid, and lends itself to automation
XX
SQ Sequence 5819 BP; 1569 A; 1348 C; 1318 G; 1584 T; 0 U; 0 Other;

Query Match	10.2%;	Score 283.8;	DB 3;	Length 5819;
Best Local Similarity	90.6%;	Pred. No. 1e-46;		
Matches 336;	Conservative 0;	Mismatches 32;	Indels 3;	Gaps 3;
QY	2405	AGCAGGCTCTTGAATTCGACCCTTCGAGCTTGGCAATCCGGTACTGTTGGTAAATGG	2464	
Db	242	AGCTGGCTGCTTAGTCAGATCTAAGTAAAGCTTGGCAATCCGGTACTGTTGGTAAATGG	301	
QY	2465	AAGACGCCAAAAACATAAGAAAGGCCCGCGCCCATTCATCTCTAGAGGATGGAAACCG	2524	
Db	302	AAGACGCCAAAAACATAAGAAAGGCCCGCGCCCATTCATCTCTAGAGGATGGAAACCG	361	
QY	2525	CTGGAGAGCAACTGCATTAAGGCTAGAGAGATACGGCCCTGGTCTCTGGAAACAATTCGCTT	2584	
Db	362	CTGGAGAGCAACTGCATTAAGGCTATGAAGAGATACGGCCCTGGTCTCTGGAAACAATTCGCTT	421	
QY	2585	TTACAGATGCACATATCGAGGTGAACATCACGTTTCGGCGAATACTTCGAATATGTCGTTT	2644	
Db	422	TTACAGATGCACATATCGAGGTGAACATCACGTTTCGGCGAATACTTCGAATATGTCG-TT	480	
QY	2645	CGTTTGGCAGAGCTATGAAACGATATGGCTGAATACAAAATCACAGAAATCGTCGTATGC	2704	
Db	481	CGTTTGGCAGAGCTATGAAACGATATGGCTGAATACAAAATCACAGAAATCGTCGTATGC	540	
QY	2705	AGTGAACAATCTCTTTTCAATCTTTTATGCGGTGTTGGGGCCGTTATTTTATCCGGAGTTG	2764	
Db	541	AGTGAACAATCTCT-TTCAATCTTTATGCGGTGTTGGGGCCGTTATTTTAT-CCGAGTTG	598	
QY	2765	CAGTTGCCGCC	2775	
b	599	CAGTTGCCGCC	609	

RESULT 8	
AAT85710	
ID	AAT85710 standard; DNA; 6092 BP.
XX	
XX	AAT85710;
XX	
XX	
XX	26-FEB-1998 (first entry)
XX	
XX	
DE	Human apolipoprotein DNA fragment from plasmid pGL2(apo AI-491)luc.
KW	
KW	Drug response element; apolipoprotein AI; apo AI gene; human;
KW	drug screening; atherosclerotic plaque; ss.
XX	
XX	
OS	Synthetic.

OS	Homo sapiens.
XX	
FN	CA2159532-A.
XX	
PD	30-MAR-1997.
XX	
PF	29-SEP-1995; 95CA-02159532.
XX	
PR	29-SEP-1995; 95CA-02159532.
XX	
PA	(TOOH) UNIV QUEBENS KINGSTON.
XX	
PI	Tam S, Zhang X;
XX	
DR	WPI; 1997-333160/31.
XX	
PT	DNA construct containing drug-responsive element of apo.lipo:protein AI
PT	gene - for transforming cells to screen for drugs that increase
PT	expression of apo AI gene.
XX	
PS	Example 1; Page 1-6 (of the sequence listing); 81pp; English.
XX	
CC	A new DNA construct has been developed which comprises the following
CC	elements functionally joined together in the 5' to 3' direction of
CC	transcription: (a) at least one copy of a drug-responsive element (DRE);
CC	(b) a heterologous promoter; (c) a reporter gene; and (d) an untranslated
CC	region including a polyadenylation signal. The present sequence
CC	represents a fragment of the human apolipoprotein gene from plasmid
CC	pGL2(apo AI-491)luc. The construct is useful for screening for drugs that
CC	increase expression of a gene for apolipoprotein AI (apo AI). If
CC	expression of the apo AI gene in humans could be increased, it would
CC	likely protect against deposition of atherosclerotic plaques
XX	
SQ	Sequence 6092 BP; 1627 A; 1424 C; 1416 G; 1625 T; 0 U; 0 Other;
Query Match	10.2%; Score 283.8; DB 2; Length 6092;
Best Local Similarity	90.6%; Pred. No. 1e-46;
Matches 336; Conservative	0; Mismatches 32; Indels 3; Gaps 3;
QY	2405 AGCAGGCTCCTTGAATTTCGACCCCTTCGAGCTTGCGATCCGGTACTGTTGGTAAATGG 2464
DB	515 AGCTGGCTGTAGTCGAGATCTAAGTAGCTTGGCATTCGGTACTGTTGGTAAATGG 574
QY	2465 AAGACGCCAAAAACATAAGAAAAGGCCCGCGCCATTCTATCTCTAGAGGATGGAAACCG 2524
DB	575 AAGACGCCAAAAACATAAGAAAAGGCCCGCGCCATTCTATCTCTAGAGGATGGAAACCG 634
QY	2525 CTGAGAGCCAACCTGCATTAAGGCTAGAAGAGATACGCCCTGGTTCCTGGAACAAATTCGCTT 2584
DB	635 CTGAGAGCCAACCTGCATTAAGGCTATGAAGAGATACGCCCTGGTTCCTGGAACAAATTCGCTT 694
QY	2585 TTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATCTCCGTTTT 2644
DB	695 TTACAGATGCACATATCGAGGTGAACATCACGTTTCGCGGAATACTTCGAAATCTCCG-TT 753
QY	2645 CGGTTGGCGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGC 2704
DB	754 CGGTTGGCGAAGCTATGAAACGATATGGGCTGAATACAAATCACAGAATCGTCGTATGC 813
QY	2705 AGTGAACAATCTCTTTCAATCTTTTATGCGGHTTGGGCCCGCTTATTTATCCGAGTTG 2764
DB	814 AGTGAACAATCTC-TTCAATCTTTTATGCGGHTTGGGCCCGCTTATTTAT-CCGAGTTG 871
QY	2765 CAGTTGCGCCC 2775
DB	872 CAGTTGCGCCC 882
RESULT 9	
AAK08773	
ID	AAK08773 standard; DNA; 6092 BP.
XX	
AC	AAK08773;

XX 27-SEP-1999 (first entry)
 XX pGL2(apo AI-491) luc plasmid construct comprising ARE/luciferase.
 DE Antioxidant responsive element; ARE; low density lipoprotein; LDL;
 XX high density lipoprotein; HDL; apolipoprotein; apo AI; atherosclerosis;
 KW heart disease; transcription; ss.
 XX Synthetic.
 OS CA2238662-A.
 XX 23-NOV-1998.
 XX 22-MAY-1998; 98CA-02238662.
 XX 23-MAY-1997; 97US-00862431.
 PR (TOOH) UNIV QUEENS KINGSTON.
 PA Tam S;
 PI WPI; 1999-229918/20.
 XX New Antioxidant Response Element (ARE), useful for identifying drugs and
 PT transcription factors for increasing transcription of mRNA, useful for
 PT treatment of atherosclerosis.
 XX Example; Page 52-55; 115pp; English.
 XX DNA constructs comprising antioxidant responsive elements (AREs) are
 CC useful for screening for compounds and transcription factors that bind to
 CC the ARE and increase transcription levels of a mRNA regulated by an ARE.
 CC AREs may also be useful as a reagent for purification of a compound
 CC (preferably a transcription factor) with which it interacts. High Density
 CC Lipoprotein (HDL) has antioxidative activity and protects against oxidized
 CC low-density lipoprotein (LDL) which has a role in the etiology of
 CC atherosclerosis. Apolipoprotein (apo) AI is a major component of HDL, and
 CC is believed to promote the process of reverse cholesterol transport. The
 CC transcription of apo AI is effected by cis- and trans-acting factors (i.e
 CC an ARE) UV cross-linking studies using an apoAI-ARE probe isolated two
 CC polypeptides of 100 and 115 kDa. These compounds are useful for treatment
 CC of a human or animal with atherosclerosis. ARE's can also be used in DNA
 CC constructs when operably linked to heterologous protein coding sequences
 CC to effect the transcription of those heterologous sequences. A Genelight
 CC vector (pGL2-B) comprising a 491 bp fragment of the apo AI from -491 to
 CC +1 inserted into the XhoI site of the vector and upstream of the
 CC luciferase gene was designated pGL2(apoAI-491)luc. Plasmid pGL2(apoAI-
 CC 250)luc was constructed by releasing a DNA fragment (-491 to -251 of the
 CC apo AI promoter) from pGL2(apoAI-491)luc. A new plasmid pGL2(apoAI-250
 CC mutant ARE) was prepared by PCR using four primers (AAx08785-88). Two
 CC primers (AAx08785, AAx08786) were complementary to the non-coding strand
 CC of DNA, the remaining two primers (AAx08787, AAx08788) were complementary
 CC to the coding strand of DNA. The plasmid pGL2(apoAI250)luc was used as a
 CC template and the DNA fragment was purified and then cloned into the SmaI
 CC and HindIII sites of pGL2-B to generate pGL2(apoAI-250 mutant ARE). A
 CC series of pGL2-Promoter vectors comprising apoAI-ARE, GST-ARE and mutated
 CC ARE were also constructed. See AAx08773-X08779
 XX Sequence 6092 BP; 1626 A; 1424 C; 1417 G; 1625 T; 0 U; 0 Other;
 SQ
 Query Match 10.2%; Score 283.8; DB 2; Length 6092;
 Best Local Similarity 90.6%; Pred. No. 1e-46;
 Matches 336; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
 XX 2405 AGCAGGCTCCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTGTGTGTAATAATCGG 2464
 DB 515 AGCTGGCTGCTTAGTGCAGATCTAAGTAGCTTGGCATTCGGTACTGTGTGTAATAATCGG 574
 XX 2465 AGACGCCCAAAACATAAAGAAAGGCCCGCGCCCATTTCTATCTCTAGAGGATGGAACCG 2524
 DB 575 AAGACGCCCAAAACATAAAGAAAGGCCCGCGCCCATTTCTATCTCTAGAGGATGGAACCG 634

QY 2525 CTGGAGAGCCAACTGCATAGGCTAGAAAGATACGCCCTCGTTCTCTGGAACAATTGCTT 2584
 DB 635 CTGGAGAGCAACTGCATAGGCTATGAAGAGATAGCCCTGGTTCTTGGACAATTGCTT 694
 QY 2585 TTACAGATGCACATATCGAGGTGAACATCAGCTTCGCGGAATACATTCGAAATGTCGGTTT 2644
 DB 695 TTACAGATGCACATATCGAGGTGAACATCAGCTACGCGGGAATACATTCGAAATGTCGG-TT 753
 QY 2645 CGTTTGGCAGAAGCTATGGAACGATATGGCTGATACAAATACAGAAATCGTCTATGC 2704
 DB 754 CGTTTGGCAGAAGCTATGGAACGATATGGCTGATACAAATACAGAAATCGTCTATGC 813
 QY 2705 AGTGAATACTCTCTTTCAATTTCTTTATCCGCTGTTGGCCCGCTTATTTATCCGAGTTG 2764
 DB 814 AGTGAATACTCTC-TTCAATTTCTTTATCCGCTGTTGGCCCGCTTATTTAT-CGAGTTG 871
 QY 2765 CAGTTGCCGCC 2775
 DB 872 CAGTTGCCGCC 882
 RESULT 10
 AAZ38350
 ID AAZ38350 standard; DNA; 6092 BP.
 XX AC AAZ38350;
 XX DT 15-FEB-2000 (first entry)
 XX DE Plasmid pGL2 (apoAI-491)luc.
 KW Apolipoprotein AI; apo AI; low density lipoprotein; LDL; cholesterol;
 KW coronary artery disease; promoter; flanking region; plasmid; DRE;
 KW drug responsive element; reporter gene; drug screening; candidate;
 KW expression; identification; rapid; automation; circular; cyclic; ds.
 OS Synthetic.
 XX US5994061-A.
 PN 30-NOV-1999.
 XX 29-SEP-1995; 95US-00536559.
 XX 29-SEP-1995; 95US-00536559.
 PR (TOOH) UNIV QUEENS KINGSTON.
 PA Tam S, Zhang X;
 PI WPI; 2000-038251/03.
 DR Screening for drugs that increase expression of the apolipoprotein AI
 XX gene, which may then be useful for treating coronary artery disease.
 XX Example 1; Col 23-30; 38pp; English.
 XX This sequence represents plasmid pGL2 (apoAI-491)luc. This plasmid
 CC comprises a 491 bp portion of the 5' flanking region of the human
 CC apolipoprotein AI (apo AI) gene inserted upstream of the luciferase
 CC reporter gene in the pGL2-Basic vector. The pGL2-Basic vector also
 CC comprises a functional polyadenylation sequence downstream of the
 CC reporter gene. This plasmid was used as a prototype for the construction
 CC of plasmid pGL2 (apoAI-250)luc (AAZ38351), pGL2 (apoAI-DRE)TK/luc
 CC (AAZ38352) and pGL2 (4xapoAI-DRE)TK/luc (AAZ38353). pGL2 (apoAI-
 CC DRE)TK/luc contains a fragment of the apoAI 5' flanking sequence (bases -
 CC 79 to -44 from the transcriptional start site) which contains two drug
 CC responsive elements (DREs, AAZ38354-238356) and an Hsv thymidine kinase
 CC (TK) promoter. pGL2 (4xapoAI-DRE)TK/luc is identical, except that it
 CC contains 4 copies of the apoAI flanking sequence fragment. Plasmids pGL2
 CC (apoAI-DRE)TK/luc and pGL2 (4xapoAI-DRE)TK/luc are used in a novel method
 CC for screening for a drug that increases expression of a gene for


```
FT      /*tag= b
FT      /note= "Luciferase gene"
FT      3701..4546
FT      /*tag= C
FT      /notes= "part of SV40"
FT      4603..6783
FT      /*tag= d
FT      /note= "EBV-oriP"
FT      9044..7159
FT      /*tag= e
FT      /notes= "hygromycin resistance gene"
FT      10994..10144
FT      /*tag= f
FT      /note= "ampicillin resistance gene"
XX
XX WO9304179-A1.
XX
XX 04-MAR-1993.
XX
XX 16-AUG-1991; 91WO-DE000652.
XX
XX 16-AUG-1991; 91WO-DE000652.
XX
XX (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
XX Bornkamm G, Polack A;
XX
XX WPI; 1993-094009/11.
XX
XX New DNA construct for detection of tumour promoters - comprises control
XX strongly-induced Epstein-Barr virus gene, and reporter gene e.g.
XX chloramphenicol acetyl transferase.
XX
XX Disclosure; Page 15; 32pp; German.
XX
```

```
CC The sequence is that of the vector pHEBO-DR-LUC contg. a control region
CC for activating a partic. strongly induced gene of the Epstein-Barr virus
CC (EBV) - the DR gene, and a reporter gene (firefly luciferase [LUC]). It
CC can be used in an in vitro test for the detection of tumour promoters
CC using EBV-contg. cells which include the vector. The in vitro test is a
CC simple and rapid alternative to animal tests and is suitable for large
CC scale screening and provides quantitative results. (Updated on 25-MAR-
CC 2003 to correct PN field.)
XX
XX Sequence 11616 BP; 3013 A; 2935 C; 2837 G; 2831 T; 0 U; 0 Other;
XX
XX Query Match 10.1%; Score 280.2; DB 2; Length 11616;
XX Best Local Similarity 93.9%; Pred. No. 5.9e-46;
XX Matches 324; Conservative 0; Mismatches 18; Indels 3; Gaps 3;
XX
QY 2431 CGAGCTTGGCATTCCGGTACTGTGTTAAATGGAAGAGCCCAAAACATAGAAGGCG 2490
DB |||||||
DB 1963 CAAGCTTGGCATTCCGGTACTGTGTTAAATGGAAGAGCCCAAAACATAGAAGGCG 2022
DB |||||||
QY 2491 CGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTAG 2550
DB |||||||
DB 2023 CGGCGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAAGGCTAG 2082
DB |||||||
QY 2551 AAGAGATAGCCCTGTTCTGGAACAATGCTTTTACAGATGCACATATCAGAGTGAAC 2610
DB |||||||
DB 2083 AAGAGATAGCCCTGTTCTGGAACAATGCTTTTACAGATGCACATATCAGAGTGAAC 2142
DB |||||||
QY 2611 ATCAGCTTCGCGAATCTTCGAAATGTCGGTTTGGCAGAGCTATGAACGATA 2670
DB |||||||
DB 2143 ATCAGCTAGCGGAATCTTCGAAATGTCGGTTTGGCAGAGCTATGAACGATA 2201
DB |||||||
QY 2671 TGGGCTGAATACAAATACAGAAATCGTATGCAAGTGAACCTCTCTTCAATTCCTTA 2730
DB |||||||
DB 2202 TGGGCTGAATACAAATACAGAAATCGTATGCAAGTGAACCTCTCTTCAATTCCTTA 2260
DB |||||||
QY 2731 TGGCGGTGTGGCCGCTTATTATCCGAGTTGCAGTTGCCGCC 2775
DB |||||||
DB 2261 TGGCGGTGTGGCCGCTTATTATCCGAGTTGCAGTTGCCGCC 2304
```

RESULT 13

```
AAQ67390
ID AAQ67390 standard; DNA; 11616 BP.
XX
XX AC AAQ67390;
XX
XX DT 25-MAR-2003 (revised)
XX DT 10-APR-1995 (first entry)
XX
XX DE Construct pHEBO-DR-REP for identifying protein kinase C inhibitors.
XX DE Protein kinase; construct; testing; screening; identification; tumour;
XX DE inflammation; Epstein-Barr virus; latent; latency; ss.
XX
XX OS Synthetic.
XX
XX FH Key
XX FT promoter
XX FT /*tag= a
XX FT /label= DR gene promoter.
XX FT /note= "Isolated from Epstein-Barr virus."
XX FT 1972..3689
XX FT /*tag= b
XX FT /label= luciferase.
XX FT 3701..4546
XX FT /*tag= c
XX FT /note= "Fragment of SV40."
XX FT 4979..7159
XX FT /*tag= d
XX FT /label= OriP
XX FT /note= "Isolated from Epstein-Barr virus."
XX FT 9044..7159
XX FT /*tag= e
XX FT /phenotype= "Hygromycin resistance."
XX FT 10994..10144
XX FT /*tag= f
XX FT /phenotype= "Ampicillin resistance."
XX FT /note= "Isolated from plasmid pBR322."
XX
XX DE4306547-A1.
XX
XX 08-SEP-1994.
XX
XX 03-MAR-1993; 93DE-04306547.
XX
XX 03-MAR-1993; 93DE-04306547.
XX
XX (GSFU-) GSF FORSCHUNGSZENTRUM UMWELT & GESUNDHEI.
XX
XX Bornkamm G, Polack A;
XX
XX WPI; 1994-280518/35.
XX
XX DNA construct contg. strongly induced Epstein-Barr virus promoter and
XX reporter gene - opt. also resistance genes, useful in in vitro
XX identification of protein kinase C inhibitors, potentially useful as
XX antitumour agents, etc.
XX
XX Disclosure; Page 5-9; 13pp; German.
XX
XX The tumour protein antigens of Epstein-Barr virus are induced by the
XX action of protein kinase C. The construct may be used to transfect EBV
XX infected cells and then used for the in vitro detection of protein kinase
XX C inhibitors which are potentially useful for treating tumours, acute
XX inflammation and activation of latent virus. It provides a simple,
XX quantitative method for detecting protein kinase C inhibitors compared
XX with known immunofluorescence techniques. (Updated on 25-MAR-2003 to
XX correct PN field.)
XX
XX Sequence 11616 BP; 3014 A; 2936 C; 2836 G; 2830 T; 0 U; 0 Other;
```


Query Match 10.1%; Score 279.8; DB 2; Length 5789;
Best Local Similarity 94.2%; Pred. No. 6.4e-46;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

Accession	Gene	Accession	Gene
2493	GGCGCCATTCTATCCTCTAGAGGATGGAAACCGCTGGAGAGGACCACTGCATAGGCTAGAA	2552	
240	AGCTTGGCATTCCGGTACTGTTGGTAAATGGGAAGACGCCAAACATATAAGANAGGCC	299	
Db			
Qv			

2493 GCGGCCATTCTATCCTCTAGAGGATGGAAACCGCTGGAGAGCCAACTGCATAAGGCTAGAA 2552

D_b 300 GCGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGAGCAACTGCATAGGCTATGAA 359

OV 2553 GAGATACGCCCTGGTTCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACAT 2612

2

Db 360 GAGATACGCCCTGGTTCCCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACAT 419

2613 CACGTTCCGGGAA^{TA}CTTCGAAATGTCCGTTTCGGTTGGCAGAAAGCTATGAAACGATATG 2672

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Db 420 CACGTACGCGGAATACTTCGAAATGTCCG-TTCGGTTGCCAGAGCTATGAAACGATATG 478

2673 GGGTGAATACAAATCACAGAAATCGTCGTATGCAGTGAAAACCTCTTTCAATTCTTTATG 2732

D_b 479 GGC TGA TACA AATC ACAGA ATCG TCG TATG CAGT GAA AACT CTC - TTCA ATTCT TATG 537

2775

2/33 2/33

Db 538 CCGGTGTTGGCGCGTTATTAT-CGGAGTGCAGTTGCGCCC 579

RESULT 15

AX08778

ID AAX08778 standard; DNA; 5791 BP.

XX AC AAX08778:

THE

DT 27-SEP-1999 (first entry)

XX DE pGL3 (CST-APF) SV40/luc plasmid construct comprising APF/luciferase.

[illegible]

KW Antioxidant responsive element; ARE; low density lipoprotein; LDL;

KW high density lipoprotein; apo AI; apolipoprotein; HDL; atherosclerosis;
KW heart disease; transcription; ap

XX
XX
XX

heart disease; transportation, see.

OS Synthetic.

XX
DNY
C C C C C C C C C C

FN
CAZ23000Z-A.
XX

PD 23-NOV-1998.

XX	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
----	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----

PF 22-MAY-1998; 98CA-0223866Z.

Search completed: August 1, 2005, 16:20:42
Job time : 1441 secs

```
XX 23-MAY-1997; 97US-00862431.
XX (TOOH ) UNIV QUEENS KINGSTON.
XX PI
XX Tam S;
XX WPI; 1999-229918/20.
XX
XX New Antioxidant Response Element (ARE), useful for identifying drugs and
XX transcription factors for increasing transcription of mRNA, useful for
XX treatment of atherosclerosis.
XX
XX Example; Page 70-74; 115pp; English.
XX
XX DNA constructs comprising antioxidant responsive elements (AREs) are
XX useful for screening for compounds and transcription factors that bind to
XX the ARE and increase transcription levels of a mRNA regulated by an ARE.
XX AREs may also be useful as a reagent for purification of a compound
XX (preferably a transcription factor) with which it interacts. High Density
XX Lipoprotein (HDL) has antioxidant activity and protects against oxidized
XX low-density lipoprotein (LDL) which has a role in the etiology of
XX atherosclerosis. Apolipoprotein (apo) AI is a major component of HDL, and
XX is believed to promote the process of reverse cholesterol transport. The
XX transcription of apo AI is effected by cis- and trans-acting factors (i.e
XX an ARE) UV cross-linking studies using an apoAI-ARE probe isolated two
XX polypeptides of 100 and 115 kDa. These compounds are useful for treatment
XX of a human or animal with atherosclerosis. ARE's can also be used in DNA
XX constructs when operably linked to heterologous protein coding sequences
XX to effect the transcription of those heterologous sequences. A Genelight
XX vector (pGL2-B) comprising a 491 bp fragment of the apo AI from -491 to
XX +1 inserted into the XhoI site of the vector and upstream of the
XX luciferase gene was designated pGL2(apoAI-491)luc. Plasmid pGL2(apoAI-
XX 250)luc was constructed by releasing a DNA fragment (-491 to -251 of the
XX apo AI promoter) from pGL2(apoAI-491)luc. A new plasmid pGL2(apoAI-250
XX mutant ARE) was prepared by PCR using four primers (AAX08785-88). Two
XX primers (AAX08785, AAX08786) were complementary to the non-coding strand
XX of DNA, the remaining two primers (AAX08787, AAX08788) were complementary
XX to the coding strand of DNA. The plasmid pGL2(apoAI-250)luc was used as a
XX template and the DNA fragment was purified and then cloned into the SmaI
XX and HindIII sites of pGL2-B to generate pGL2(apoAI-250 mutant ARE). A
XX series of pGL2-Promoter vectors comprising apoAI-ARE, GST-ARE and mutated
XX ARE were also constructed. See AAX08773-X08779
XX
XX Sequence 5791 BP; 1564 A; 1328 C; 1306 G; 1593 T; 0 U; 0 Other;
XX
Query Match 10.1%; Score 279.8; DB 2; Length 5791;
Best Local Similarity 94.2%; Pred. No. 6.4e-46;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 2433 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGACGCCAAACATAAAGAAAGGCC 2492
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
242 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGACGCCAAACATAAAGAAAGGCC 301
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2493 GGCGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAAGGCTAGAA 2552
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
302 GGCGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAAGGCTATGAA 161
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2553 GAGATACGGCCCTGGTTCCTGGAAACAATTTGTTTACAGATGCACATATCGAGGTGAACAT 2612
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
362 GAGATACGGCCCTGGTTCCTGGAAACAATTTGTTTACAGATGCACATATCGAGGTGAACAT 421
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2613 CAGGTTCCGGGATACCTCGAATGTCCTGTTGGTTGGCAGAGCTATGAACGATATG 2672
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
422 CAGGTACGGGGAATACCTCGAATGTCCTGTTGGTTGGCAGAGCTATGAACGATATG 480
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2673 GGCTGAATACAAATCAGAAATCGTGTATGCAGTGAACCTCTCTTTCAATTTCTTTATG 2732
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
481 GGCTGAATACAAATCAGAAATCGTGTATGCAGTGAACCTCTCTTTCAATTTCTTTATG 539
QY |||||||||||||||||||||||||||||||||||||||||||||||||||||||
2733 CCGGTGTGGGCCCGTATTTATTCGGAGTTGCGAGTTGCGGCC 2775
Db |||||||||||||||||||||||||||||||||||||||||||||||||||||||
540 CCGGTGTGGGCCCGTATTTATTCGGAGTTGCGAGTTGCGGCC 581
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: August 1, 2005, 07:20:48 ; Search time 1706 Seconds
(without alignments)

10562.157 Million cell updates/sec

Title: US-09-068-751-1

Perfect score: 2784

Sequence: 1 Gaagtgaatctgataatt.....cagttgcgcgcgcgaaca 2784

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7287783 seqs, 3236178273 residues

Total number of hits satisfying chosen parameters: 14575566

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10G_PUBCOMB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US10H_PUBCOMB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US10I_PUBCOMB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 23: /cgn2_6/ptodata/2/pubpna/US11A_PUBCOMB.seq.*
- 24: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 25: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 26: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c					
1	393.6	14.1	493631	13	US-10-087-192-205 Sequence 205, Appl
2	284.6	10.2	4672	14	US-10-059-561-5 Sequence 5, Appli
3	279.8	10.1	5849	20	US-10-677-777-1 Sequence 1, Appli
4	279.8	10.1	8937	9	US-09-872-733-8 Sequence 8, Appli
5	279.8	10.1	8937	9	US-09-872-733-9 Sequence 9, Appli
6	279.8	10.1	8937	14	US-10-263-020-8 Sequence 8, Appli
7	279.8	10.1	8937	14	US-10-263-020-9 Sequence 9, Appli

8	279.8	10.1	8937	18	US-10-644-027-8	Sequence 8, Appli	
9	279.8	10.1	8937	18	US-10-644-027-9	Sequence 9, Appli	
10	276.8	9.9	5620	9	US-09-918-029-21	Sequence 21, Appl	
11	276.8	9.9	5620	9	US-09-912-552-21	Sequence 21, Appl	
12	276.8	9.9	5620	13	US-10-038-271-21	Sequence 21, Appl	
13	276.8	9.9	5620	13	US-10-125-751-21	Sequence 21, Appl	
14	276.8	9.9	5620	15	US-10-219-414-21	Sequence 21, Appl	
15	276.8	9.9	5620	20	US-10-850-140-21	Sequence 21, Appl	
16	274.6	9.9	2502	16	US-10-220-262-8	Sequence 8, Appli	
c	17	268.4	9.6	720	13	US-10-101-487-74	Sequence 74, Appl
18	268.4	9.6	720	13	US-10-101-487-76	Sequence 74, Appl	
c	19	268.4	9.6	720	21	US-10-939-988-74	Sequence 74, Appl
20	268.4	9.6	720	21	US-10-939-988-76	Sequence 76, Appl	
21	254.6	9.1	1798	9	US-09-956-988A-16	Sequence 16, Appl	
22	254.6	9.1	4824	21	US-10-169-050-52	Sequence 52, Appl	
23	254.6	9.1	5376	20	US-10-492-396-7	Sequence 7, Appli	
24	254.6	9.1	5960	10	US-09-798-883B-14	Sequence 14, Appl	
25	254.6	9.1	5960	10	US-09-326-885-14	Sequence 14, Appl	
26	254.2	9.1	5924	20	US-10-790-455-3	Sequence 3, Appli	
27	254.2	9.1	5924	20	US-10-811-136B-3	Sequence 3, Appli	
28	254.2	9.1	5924	21	US-10-940-315-3	Sequence 3, Appli	
29	254.2	9.1	5924	21	US-10-950-050-3	Sequence 3, Appli	
30	254.2	9.1	5982	20	US-10-790-455-2	Sequence 2, Appli	
31	254.2	9.1	5982	20	US-10-811-136B-2	Sequence 2, Appli	
32	254.2	9.1	5982	21	US-10-940-315-2	Sequence 2, Appli	
33	254.2	9.1	5982	21	US-10-950-050-2	Sequence 2, Appli	
34	253.8	9.1	3519	15	US-10-122-706-3	Sequence 3, Appli	
35	253.8	9.1	3519	22	US-10-494-073-3	Sequence 3, Appli	
36	253.8	9.1	5777	19	US-10-467-781A-3	Sequence 3, Appli	
37	253.8	9.1	36083	21	US-10-944-496-21	Sequence 21, Appl	
38	253.2	9.1	5742	19	US-10-688-299-47	Sequence 47, Appl	
39	253.2	9.1	6525	14	US-10-055-794-3	Sequence 3, Appli	
40	253.2	9.1	7487	14	US-10-055-794-4	Sequence 4, Appli	
c	41	253	9.1	2477	16	US-10-087-167-103	Sequence 103, App
42	253	9.1	5143	21	US-10-169-050-49	Sequence 49, Appl	
43	253	9.1	5662	21	US-10-169-050-50	Sequence 50, Appl	
44	253	9.1	6310	21	US-10-169-050-48	Sequence 48, Appl	
45	253	9.1	15185	15	US-10-121-960C-17	Sequence 17, Appl	

ALIGNMENTS

RESULT 1
US-10-087-192-205/c
; Sequence 205, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 205
; LENGTH: 493631
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(493631)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-205

Query Match 14.1%; Score 393.6; DB 13; Length 493631;
Best Local Similarity 41.7%; Pred. No. 9.5e-104;

Db 553 C-TTCAATTCTTTATGCGGGTGTGGCGGCTTATTTAT-CGGAGTTGCAGTTGCGGCC 609

RESULT 3
US-10-677-777-1
; Sequence 1, Application US/1067777
; Publication No. US20040235157A1
; GENERAL INFORMATION:
; APPLICANT: TOVEY, Michael
; TITLE OF INVENTION: GENE REPORTER ASSAY, KIT, AND CELLS FOR DETERMINING THE PRESENCE
; TITLE OF INVENTION: AND/OR THE LEVEL OF A MOLECULE THAT ACTIVATES SIGNAL TRANSDUCTION
; TITLE OF INVENTION: ACTIVITY OF A CELL SURFACE PROTEIN
; FILE REFERENCE: TOVEY=5A
; CURRENT APPLICATION NUMBER: US/10/677,777
; CURRENT FILING DATE: 2003-10-03
; PRIOR APPLICATION NUMBER: 60/415,818
; PRIOR FILING DATE: 2002-10-04
; NUMBER OF SEQ ID NOS: 7
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 5849
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic
; NAME/KEY: CDS
; LOCATION: (328)..(1980)
US-10-677-777-1

Query Match 10.1%; Score 279.8; DB 20; Length 5849;
Best Local Similarity 94.2%; Pred. No. 4.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

Qy 2433 AGCTTGGCATTCCGGTACTGTGTAAATGGAAGACGCCCAAAACATTAAGAAAGGCC 2492
Db 300 AGCTTGGCATTCCGGTACTGTGTAAATGGAAGACGCCCAAAACATTAAGAAAGGCC 359
Qy 2493 GGCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGGCCAACTGCATTAAGGCTAGAA 2552
Db 360 GGCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGGCCAACTGCATTAAGGCTAGAA 419
Qy 2553 GAGATACGCCCTGTCTCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 420 GAGATACGCCCTGTCTCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACAT 479
Qy 2613 CACGTTCCGCGAATCTCGAATGTCGGTTCCGTTGGCAGAGCTATGAAACGATATG 2672
Db 480 CACGTTCCGCGAATCTCGAATGTCGGTTCCGTTGGCAGAGCTATGAAACGATATG 538
Qy 2673 GGCTGAATACAAATACAGAAATCGTCTATGCAAGTAAACTCTCTTCAATTCTTTATG 2732
Db 539 GGCTGAATACAAATACAGAAATCGTCTATGCAAGTAAACTCTCTTCAATTCTTTATG 597
Qy 2733 CCGGTGTGGCGCGTATTATTCGGAGTTGCAGTTGCGGCC 2775
Db 598 CCGGTGTGGCGCGTATTATTCGGAGTTGCAGTTGCGGCC 639

RESULT 4
US-09-872-733-8
; Sequence 8, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22

Qy 2433 AGCTTGGCATTCCGGTACTGTGTAAATGGAAGACGCCCAAAACATTAAGAAAGGCC 2492

; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBcWcNlucI
US-09-872-733-8

Query Match 10.1%; Score 279.8; DB 9; Length 8937;
Best Local Similarity 94.2%; Pred. No. 5.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

Qy 2433 AGCTTGGCATTCCGGTACTGTGTAAATGGAAGACGCCCAAAACATTAAGAAAGGCC 2492
Db 1867 AGCTTGGCATTCCGGTACTGTGTAAATGGAAGACGCCCAAAACATTAAGAAAGGCC 1926
Qy 2493 GGCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGGCCAACTGCATTAAGGCTAGAA 2552
Db 1927 GGCGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGGCCAACTGCATTAAGGCTAGAA 1986
Qy 2553 GAGATACGCCCTGTCTCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 1987 GAGATACGCCCTGTCTCTGGAACAATTGCTTTTACAGATGCACATATCGAGGTGAACAT 2046
Qy 2613 CACGTTCCGCGAATCTCGAATGTCGGTTCCGTTGGCAGAGCTATGAAACGATATG 2672
Db 2047 CACGTTCCGCGAATCTCGAATGTCGGTTCCGTTGGCAGAGCTATGAAACGATATG 2105
Qy 2673 GGCTGAATACAAATACAGAAATCGTCTATGCAAGTAAACTCTCTTCAATTCTTTATG 2732
Db 2106 GGCTGAATACAAATACAGAAATCGTCTATGCAAGTAAACTCTCTTCAATTCTTTATG 2164
Qy 2733 CCGGTGTGGCGCGTATTATTCGGAGTTGCAGTTGCGGCC 2775
Db 2165 CCGGTGTGGCGCGTATTATTCGGAGTTGCAGTTGCGGCC 2206

RESULT 5
US-09-872-733-9
; Sequence 9, Application US/09872733
; Patent No. US20010036655A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/09/872,733
; CURRENT FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: from transfer construct pmBcWcNlucI
US-09-872-733-9

Query Match 10.1%; Score 279.8; DB 9; Length 8937;
Best Local Similarity 94.2%; Pred. No. 5.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

Qy 2433 AGCTTGGCATTCCGGTACTGTGTAAATGGAAGACGCCCAAAACATTAAGAAAGGCC 2492

Db 1867 AGCTTGGCAATCCGGTACTGTTGGTAAATGAAGACGCCAAAAACATAAAGAAAGGCC 1926
QY 2493 GCGGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCTAAGGCTAGAA 2552
Db 1927 GCGGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCTAAGGCTATGAA 1986
QY 2553 GAGATACGCCCTGGTTTCCTGGAACTATGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 1987 GAGATACGCCCTGGTTTCCTGGAACTATGCTTTTACAGATGCACATATCGAGGTGAACAT 2046
QY 2613 CACGTTCCGGGAATATCTCGAAATGTCGTTTCGGTTGGCAGAACATATGAACGATATG 2672
Db 2047 CACGTACGGGAATATCTCGAAATGTCGTTTCGGTTGGCAGAACATATGAACGATATG 2105
QY 2673 GCGTGAATACAATCACAGAATCGTGTATGAGTGAGTGAACACTCTCTTCAATTCTTTATG 2732
Db 2106 GCGTGAATACAATCACAGAATCGTGTATGAGTGAGTGAACACTCTC-TTCAATTCTTTATG 2164
QY 2733 CCGGTGTTGGGCCGCTTATTTATCCGAGTTGCAGTTGCGGCC 2775
Db 2165 CCGGTGTTGGGCCGCTTATTTAT-CGGAGTTGCAGTTGCGGCC 2206

RESULT 6

US-10-263-020-8
; Sequence 8, Application US/10263020
; Publication No. US20030049229A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/10/263,020
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US/09/872,733
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBCwCNIuci
US-10-263-020-8

Query Match 10.1%; Score 279.8; DB 14; Length 8937;
Best Local Similarity 94.2%; Pred. No. 5.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 2433 AGCTTGGCAATCCGGTACTGTTGGTAAATGAAGACGCCAAAAACATAAAGAAAGGCC 2492
Db 1867 AGCTTGGCAATCCGGTACTGTTGGTAAATGAAGACGCCAAAAACATAAAGAAAGGCC 1926
QY 2493 GCGGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 2552
Db 1927 GCGGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTATGAA 1986
QY 2553 GAGATACGCCCTGGTTTCCTGGAACTATGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 1987 GAGATACGCCCTGGTTTCCTGGAACTATGCTTTTACAGATGCACATATCGAGGTGAACAT 2046
QY 2613 CACGTTCCGGGAATATCTCGAAATGTCGTTTCGGTTGGCAGAACATATGAACGATATG 2672
Db 2047 CACGTACGGGAATATCTCGAAATGTCGTTTCGGTTGGCAGAACATATGAACGATATG 2105
QY 2673 GCGTGAATACAATCACAGAATCGTGTATGAGTGAGTGAACACTCTCTTCAATTCTTTATG 2732

Db 2106 GGCTGAATACAATCACAGAATCGTGTATGAGTGAGTGAACACTCTC-TTCAATTCTTTATG 2164
QY 2733 CCGGTGTTGGGCCGCTTATTTATCGGAGTTGCAGTTGCGGCC 2775
Db 2165 CCGGTGTTGGGCCGCTTATTTAT-CGGAGTTGCAGTTGCGGCC 2206

RESULT 7

US-10-263-020-9
; Sequence 9, Application US/10263020
; Publication No. US20030049229A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/10/263,020
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: US/09/872,733
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBCwCNIuci
US-10-263-020-9

Query Match 10.1%; Score 279.8; DB 14; Length 8937;
Best Local Similarity 94.2%; Pred. No. 5.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;
QY 2433 AGCTTGGCAATCCGGTACTGTTGGTAAATGAAGACGCCAAAAACATAAAGAAAGGCC 2492
Db 1867 AGCTTGGCAATCCGGTACTGTTGGTAAATGAAGACGCCAAAAACATAAAGAAAGGCC 1926
QY 2493 GCGGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTAGAA 2552
Db 1927 GCGGCCATTCTATCCTCTAGAGATGGAACCGCTGGAGAGCAACTGCATAGGCTATGAA 1986
QY 2553 GAGATACGCCCTGGTTTCCTGGAACTATGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 1987 GAGATACGCCCTGGTTTCCTGGAACTATGCTTTTACAGATGCACATATCGAGGTGAACAT 2046
QY 2613 CACGTTCCGGGAATATCTCGAAATGTCGTTTCGGTTGGCAGAACATATGAACGATATG 2672
Db 2047 CACGTACGGGAATATCTCGAAATGTCGTTTCGGTTGGCAGAACATATGAACGATATG 2105
QY 2673 GCGTGAATACAATCACAGAATCGTGTATGAGTGAGTGAACACTCTCTTCAATTCTTTATG 2732
Db 2106 GCGTGAATACAATCACAGAATCGTGTATGAGTGAGTGAACACTCTC-TTCAATTCTTTATG 2164
QY 2733 CCGGTGTTGGGCCGCTTATTTATCCGAGTTGCAGTTGCGGCC 2775
Db 2165 CCGGTGTTGGGCCGCTTATTTAT-CGGAGTTGCAGTTGCGGCC 2206

RESULT 8

US-10-644-027-8
; Sequence 8, Application US/10644027
; Publication No. US2004007757A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL, SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/10/644,027
; CURRENT FILING DATE: 2004-07-27
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBCwCNIuci
US-10-644-027-8

```
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/10/644,027
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/872,733A
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBwCnlucl
US-10-644-027-8

Query Match      10.1%; Score 279.8; DB 18; Length 8937;
Best Local Similarity 94.2%; Pred. No. 5.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2433 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATATAAGAAAGGCC 2492
Db 1867 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATATAAGAAAGGCC 1926
QY 2493 GCGGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAAGCGCTAGAA 2552
Db 1927 GCGGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAAGCGCTAGAA 1986
QY 2553 GAGATACGCCCTGTTCCCTGGAACAATGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 1987 GAGATACGCCCTGTTCCCTGGAACAATGCTTTTACAGATGCACATATCGAGGTGAACAT 2046
QY 2613 CACGTTGCGGGAATACTTCCGAAATGTCGGTTTGGCAGAAGCTATGAAAACGATATG 2672
Db 2047 CACGTTGCGGGAATACTTCCGAAATGTCGGTTTGGCAGAAGCTATGAAAACGATATG 2105
QY 2673 GCGTGAATACAAATACAGAAATCGTCGTATGCAAGTGAAGAACTCTTTCAATTCTTTATG 2732
Db 2106 GCGTGAATACAAATACAGAAATCGTCGTATGCAAGTGAAGAACTCTT-TCATTCTTTATG 2164
QY 2733 CCGGTGTGGCGCGTATTATTCGGAGTTGCGAGTTGCCGCC 2775
Db 2165 CCGGTGTGGCGCGTATTATTCGGAGTTGCGAGTTGCCGCC 2206

RESULT 10
US-09-918-029-21
; Sequence 21, Application US/09918029
; Patent No. US20020102732A1
; GENERAL INFORMATION:
; APPLICANT: Fallaux, Frits J.
; APPLICANT: Hoebe, Robert C.
; APPLICANT: Bout, Abraham
; APPLICANT: Valerio, Domenico
; APPLICANT: Van der Eb, Alex J.
; TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT ADENOVIRUS TO
; TITLE OF INVENTION: BE USED
; TITLE OF INVENTION: IN GENE THERAPY
; FILE REFERENCE: 3833 2US
; CURRENT APPLICATION NUMBER: US/09/918,029
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/506,548
; PRIOR FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: PCT/NL96/00244
; PRIOR FILING DATE: 1996-06-14
; PRIOR APPLICATION NUMBER: EP 95201728.3
; PRIOR FILING DATE: 1995-06-26
; PRIOR APPLICATION NUMBER: EP 95201611.1
; PRIOR FILING DATE: 1995-06-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: Corel WordPerfect 8.0
; SEQ ID NO 21
; LENGTH: 5620
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: Ad5 left terminus
; LOCATION: 1..457
; NAME/KEY: enhancer
; LOCATION: 458..969
; NAME/KEY: exon
; LOCATION: 970..1204
; NAME/KEY: gene
; LOCATION: 1218..2987
; NAME/KEY: polyA_signal

; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/10/644,027
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/872,733A
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: DNA sequence
; OTHER INFORMATION: of transfer construct pmBwCnlucl
US-10-644-027-9

Query Match      10.1%; Score 279.8; DB 18; Length 8937;
Best Local Similarity 94.2%; Pred. No. 5.9e-71;
Matches 323; Conservative 0; Mismatches 17; Indels 3; Gaps 3;

QY 2433 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATATAAGAAAGGCC 2492
Db 1867 AGCTTGGCATTCCGGTACTGTTGGTAAATGGAAGACGCCCAAAACATATAAGAAAGGCC 1926
QY 2493 GCGGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAAGCGCTAGAA 2552
Db 1927 GCGGCCATTCTATCTCTAGAGATGGAACCGCTGGAGAGCCAACTGCATAAGCGCTAGAA 1986
QY 2553 GAGATACGCCCTGTTCCCTGGAACAATGCTTTTACAGATGCACATATCGAGGTGAACAT 2612
Db 1987 GAGATACGCCCTGTTCCCTGGAACAATGCTTTTACAGATGCACATATCGAGGTGAACAT 2046
QY 2613 CACGTTGCGGGAATACTTCCGAAATGTCGGTTTGGCAGAAGCTATGAAAACGATATG 2672
Db 2047 CACGTTGCGGGAATACTTCCGAAATGTCGGTTTGGCAGAAGCTATGAAAACGATATG 2105
QY 2673 GCGTGAATACAAATACAGAAATCGTCGTATGCAAGTGAAGAACTCTTTCAATTCTTTATG 2732
Db 2106 GCGTGAATACAAATACAGAAATCGTCGTATGCAAGTGAAGAACTCTT-TCATTCTTTATG 2164
QY 2733 CCGGTGTGGCGCGTATTATTCGGAGTTGCGAGTTGCCGCC 2775
Db 2165 CCGGTGTGGCGCGTATTATTCGGAGTTGCGAGTTGCCGCC 2206

RESULT 9
US-10-644-027-9
; Sequence 9, Application US/10644027
; Publication No. US2004007757A1
; GENERAL INFORMATION:
; APPLICANT: The Government of the United States of America, as
; TITLE OF INVENTION: MOLECULAR CLONES WITH MUTATED HIV GAG/POL, SIV GAG AND
; TITLE OF INVENTION: SIV ENV GENES
; FILE REFERENCE: 2026-4287US1 HIV GAG/POL,SIV GAG & ENV
; CURRENT APPLICATION NUMBER: US/10/644,027
; CURRENT FILING DATE: 2003-08-19
; PRIOR APPLICATION NUMBER: US/09/872,733A
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US00/34985
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/173,036
; PRIOR FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 8937
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
```

LOCATION: 3018..3131
NAME/KEY: pUC12 backbone
LOCATION: 3132..5620
NAME/KEY: Gene
LOCATION: 4337..5191
OTHER INFORMATION: Description of Artificial Sequence: Plasmid pUC12
US-09-918-029-21

Query Match 9.9%; Score 276.8; DB 9; Length 5620;
Best Local Similarity 90.4%; Pred. No. 3.7e-70;
Matches 329; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2412 TCCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTGTTGGTAAATGGAAGACGC 2471
DB 1237 TCTTGAATGTCGCTCGCAGTGACATTTAGCATTCGGGTACTGTTGGTAAATGGAAGACGC 1296
QY 2472 CAAAAACATAAAGAAAGCCGCGCCCATTCCTCTAGAGGATGGAACCGCTGGAGA 2531
DB 1297 CAAAAACATAAAGAAAGCCGCGCCCATTCCTCTAGAGGATGGAACCGCTGGAGA 1356
QY 2532 GCCAAGCTGCATAAGGCTAGAGAGATACGCCCTTCTGGAACAATTCGCTTTTACAGA 2591
DB 1357 GCNACTGCATAGGCTATGAAGAAATACGCCCTGTTCTGGAACAATTCGCTTTTACAGA 1416
QY 2592 TGCACATATCGAGGTGAACATCACGTTTCGCGAATACCTTCGAAATGTCGGTTTGG 2651
DB 1417 TGCACATATCGAGGTGAACATCACGTTTCGCGAATACCTTCGAAATGTCGGTTTGG 1475
QY 2652 CAGAAGCTATGAACCATATGGCTGAATACAAATCAGAAATCGTCGATGCAAGTGA 2711
DB 1476 CAGAAGCTATGAACCATATGGCTGAATACAAATCAGAAATCGTCGATGCAAGTGA 1535
QY 2712 ACTCTCTTTCAATTCCTTTATGCGCGTGTGGGCCGCTTATTTATCCGAGTTTGCAGTTGC 2771
DB 1536 ACTCTC-TTCAATTCCTTTATGCGCGTGTGGGCCGCTTATTTAT-CCGAGTTTGCAGTTGC 1593
QY 2772 CGCC 2775
DB 1594 GCCC 1597

RESULT 11
US-09-912-552-21
Sequence 21, Application US/09912552
Publication No. US20020187553A1
GENERAL INFORMATION:
APPLICANT: Fallaux, Frits
APPLICANT: Hoebein, Robert
APPLICANT: Bout, Abraham
APPLICANT: Valerio, Domenico
APPLICANT: van der Eb, Alex
APPLICANT: Schouten, Goevert
TITLE OF INVENTION: PACKAGING SYSTEMS
FILE REFERENCE: 2578-3935US
CURRENT APPLICATION NUMBER: US/09/912,552
CURRENT FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: US/09/356,575
PRIOR FILING DATE: 1999-07-19
PRIOR APPLICATION NUMBER: PCT/NL96/00244
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: EP 95201611.1
PRIOR FILING DATE: 1995-06-15
PRIOR APPLICATION NUMBER: EP 95201728.3
PRIOR FILING DATE: 1995-06-26
NUMBER OF SEQ ID NOS: 22
SOFTWARE: PatentIn version 3.0
SEQ ID NO 21
LENGTH: 5620
TYPE: DNA
ORGANISM: Unknown
FEATURE:
OTHER INFORMATION: Derived from Adenovirus
US-09-912-552-21

Query Match 9.9%; Score 276.8; DB 9; Length 5620;
Best Local Similarity 90.4%; Pred. No. 3.7e-70;
Matches 329; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2412 TCCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTGTTGGTAAATGGAAGACGC 2471
DB 1237 TCTTGAATGTCGCTCGCAGTGACATTTAGCATTCGGGTACTGTTGGTAAATGGAAGACGC 1296
QY 2472 CAAAAACATAAAGAAAGCCGCGCCCATTCCTCTAGAGGATGGAACCGCTGGAGA 2531
DB 1297 CAAAAACATAAAGAAAGCCGCGCCCATTCCTCTAGAGGATGGAACCGCTGGAGA 1356
QY 2532 GCCAAGCTGCATAAGGCTAGAGAGATACGCCCTTCTGGAACAATTCGCTTTTACAGA 2591
DB 1357 GCNACTGCATAGGCTATGAAGAAATACGCCCTGTTCTGGAACAATTCGCTTTTACAGA 1416
QY 2592 TGCACATATCGAGGTGAACATCACGTTTCGCGAATACCTTCGAAATGTCGGTTTGG 2651
DB 1417 TGCACATATCGAGGTGAACATCACGTTTCGCGAATACCTTCGAAATGTCGGTTTGG 1475
QY 2652 CAGAAGCTATGAACCATATGGCTGAATACAAATCAGAAATCGTCGATGCAAGTGA 2711
DB 1476 CAGAAGCTATGAACCATATGGCTGAATACAAATCAGAAATCGTCGATGCAAGTGA 1535
QY 2712 ACTCTCTTTCAATTCCTTTATGCGCGTGTGGGCCGCTTATTTATCCGAGTTTGCAGTTGC 2771
DB 1536 ACTCTC-TTCAATTCCTTTATGCGCGTGTGGGCCGCTTATTTAT-CCGAGTTTGCAGTTGC 1593
QY 2772 CGCC 2775
DB 1594 GCCC 1597
RESULT 12
US-10-038-271-21
Sequence 21, Application US/10038271
Publication No. US20020151032A1
GENERAL INFORMATION:
APPLICANT: Fallaux, Frits J.
APPLICANT: Hoebein, Robert C.
APPLICANT: Bout, Abraham
APPLICANT: Valerio, Domenico
APPLICANT: Van der Eb, Alex J.
TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT ADENOVIRUS TO BE USED IN
FILE REFERENCE: 3833.IUS
CURRENT APPLICATION NUMBER: US/10/038,271
CURRENT FILING DATE: 2001-10-23
PRIOR APPLICATION NUMBER: 09/333,820
PRIOR FILING DATE: 1999-06-15
PRIOR APPLICATION NUMBER: US 08/793,170
PRIOR FILING DATE: 1997-03-25
PRIOR APPLICATION NUMBER: PCT/NL96/00244
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: EP 95201728.3
PRIOR FILING DATE: 1995-06-26
PRIOR APPLICATION NUMBER: EP 95201611.1
PRIOR FILING DATE: 1995-06-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Corel Wordperfect 8.0
SEQ ID NO 21
LENGTH: 5620
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Ads left terminus
LOCATION: 1..457
NAME/KEY: enhancer
LOCATION: 458..969
NAME/KEY: exon
LOCATION: 970..1204
NAME/KEY: gene

FEATURE:	NAME/KEY: Ad5 left terminus	Score 276.8; DB 13; Length 5620;
LOCATION:	1..457	
FEATURE:	NAME/KEY: enhancer	
LOCATION:	458..969	
FEATURE:	NAME/KEY: exon	
LOCATION:	970..1204	
FEATURE:	NAME/KEY: Gene	
LOCATION:	1218..2987	
FEATURE:	NAME/KEY: polyA_signal	
LOCATION:	3018..3131	
FEATURE:	NAME/KEY: pUC12 backbone	
LOCATION:	3132..5620	
FEATURE:	NAME/KEY: gene	
LOCATION:	4337..5191	
OTHER INFORMATION:	Description of Artificial Sequence: Plasmid pICL	
US-10-125-751-21		
Query Match	9.9%; Score 276.8; DB 13; Length 5620;	
Best Local Similarity	90.4%; Pred. No. 3.7e-70;	
Matches 329; Conservative	0; Mismatches 32; Indels 3; Gaps 3;	
Qy	2412 TCCTTCAATTCGACCCCTTCGAGCTTGGCATTCGGTACTGTTGGTAAATGGAAGACGC	2471
Db	1237 TCTTGAATGTCGCTCGAGTGACATTAGCANTCCGGTACTGTTGGTAAATGGAAGACGC	1296
Qy	2472 CAAAAACATAAAGAAAGGCGCGGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGA	2531
Db	1297 CAAAAACATAAAGAAAGGCGCGGCCATTCTATCCTCTAGAGGATGGAACCGCTGGAGA	1356
Qy	2532 GCCAATGCATAGGCTAGAGATACGCCCTGGTTCCTGGAAACAATTGCTTTTACAGA	2591
Db	1357 GCAACTGCATAGGCTATGAAGAAATACGCCCTGGTTCCTGGAAACAATTGCTTTTACAGA	1416
Qy	2592 TGCACATATCGAGTGAACATACAGTTCGCGGAATCTTCGAAATGTCGGTTTCGGTTGG	2651
Db	1417 TGCACATATCGAGTGAACATACAGTTCGCGGAATCTTCGAAATGTCGGTTTCGGTTGG	1475
Qy	2652 CAGAAGCTATGAACGATATGGGCTCAATAACAAATCAGAAATCGTCGTATGCAGTGAAG	2711
Db	1476 CAGAAGCTATGAACGATATGGGCTCAATAACAAATCAGAAATCGTCGTATGCAGTGAAG	1535
Qy	2712 ACTCTCTTCAATCTTTATGCGCGTGGTGGGCCCGTTATTTATCGGAGTTCGAGTTGC	2771
Db	1536 ACTCTC-TTCAATCTTTATGCGCGTGGTGGGCCCGTTATTTAT-CGGAGTTCGAGTTGC	1593
Qy	2772 CGCC 2775	
Db	1594 GCCC 1597	
RESULT 14		
US-10-219-414-21		
Sequence 21, Application US/10219414		
Publication No. US20030104626A1		
GENERAL INFORMATION:		
APPLICANT: Fallaux, Frits J.		
APPLICANT: Hoeben, Robert C.		
APPLICANT: Bout, Abraham		
APPLICANT: Valerio, Domenico		
APPLICANT: Van der Eb, Alex J.		
TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT ADENOVIRUS TO		
TITLE OF INVENTION: BE USED		
TITLE OF INVENTION: IN GENE THERAPY		
FILE REFERENCE: 3833.2US		
CURRENT APPLICATION NUMBER: US/10/219,414		
CURRENT FILING DATE: 2002-08-15		

APPLICANT: Hoeben, Robert C.
APPLICANT: Bout, Abraham
APPLICANT: Valerio, Domenico
APPLICANT: Van der Eb, Alex J.
TITLE OF INVENTION: PACKAGING SYSTEMS FOR HUMAN RECOMBINANT ADENOVIRUS TO
TITLE OF INVENTION: BE USED
FILE REFERENCE: 3833.2US
CURRENT APPLICATION NUMBER: US/10/850,140
CURRENT FILING DATE: 2004-05-20
PRIOR APPLICATION NUMBER: US/09/918,029
PRIOR FILING DATE: 2001-07-30
PRIOR APPLICATION NUMBER: 09/506,548
PRIOR FILING DATE: 2000-02-16
PRIOR APPLICATION NUMBER: PCT/NL96/00244
PRIOR FILING DATE: 1996-06-14
PRIOR APPLICATION NUMBER: EP 95201728.3
PRIOR FILING DATE: 1995-06-26
PRIOR APPLICATION NUMBER: EP 95201611.1
PRIOR FILING DATE: 1995-06-15
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Corel WordPerfect 8.0
SEQ ID NO 21
LENGTH: 5620
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
NAME/KEY: Ad5 left terminus
LOCATION: 1..457
NAME/KEY: enhancer
LOCATION: 458..969
NAME/KEY: exon
LOCATION: 970..1204
NAME/KEY: gene
LOCATION: 1218..2987
NAME/KEY: polyA signal
LOCATION: 3018..3131
NAME/KEY: pUC12 backbone
LOCATION: 3132..5620
NAME/KEY: gene
LOCATION: 4337..5191
OTHER INFORMATION: Description of Artificial Sequence: Plasmid pICL
US-10-219-414-21

Query Match 9.9%; Score 276.8; DB 15; Length 5620;
Best Local Similarity 90.4%; Pred. No. 3.7e-70;
Matches 329; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2412 TCCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTGTTGGTAAATGGAAGACGC 2471
DB 1237 TCTTGAATGTCGCTCGCAGTGACATTAGCATTCGGTACTGTTGGTAAATGGAAGACGC 1296
QY 2472 CAAAAACATAAAGAGAGCCGCGCATTCCTCTAGAGGATGGAACCGCTGGAGA 2531
DB 1297 CAAAAACATAAAGAGAGCCGCGCATTCCTCTAGAGGATGGAACCGCTGGAGA 1356
QY 2532 GCCAATCGATTAAGCTTAGAAGATACGCCCTGTTCTCTGGAACAATTCGCTTTACAGA 2591
DB 1357 GCCAATCGATTAAGCTTAGAAGATACGCCCTGTTCTCTGGAACAATTCGCTTTACAGA 1416
QY 2592 TGCACATATCGAGGTGAACATCACGTTCCGCGAATACCTCGAATGTCGGTTCCGTTGG 2651
DB 1417 TGCACATATCGAGGTGAACATCACGTTCCGCGAATACCTCGAATGTCGGTTCCGTTGG 1475
QY 2652 CAGAACTATGAACCATATGGCTGAATACAAATCACAGAATCGTCGATGCAAGTAAA 2711
DB 1476 CAGAACTATGAACCATATGGCTGAATACAAATCACAGAATCGTCGATGCAAGTAAA 1535
QY 2712 ACTCTCTTTCAATCTTTTATGCGGTGTTGGGCCCGCTTTATTTATCCGAGTTGCAAGTTGC 2771
DB 1536 ACTCTC-TTCAATCTTTTATGCGGTGTTGGGCCCGCTTTATTTAT-CGGAGTTGCAAGTTGC 1593
QY 2772 CGCC 2775
DB 1594 GCCC 1597

RESULT 15

US-10-850-140-21
; Sequence 21, Application US/10850140
; Publication No. US20040228843A1
; GENERAL INFORMATION:
; APPLICANT: Fallaux, Frits J.

Query Match 9.9%; Score 276.8; DB 20; Length 5620;
Best Local Similarity 90.4%; Pred. No. 3.7e-70;
Matches 329; Conservative 0; Mismatches 32; Indels 3; Gaps 3;
QY 2412 TCCTTGAATTCGACCCCTTCGAGCTTGGCATTCGGTACTGTTGGTAAATGGAAGACGC 2471
DB 1237 TCTTGAATGTCGCTCGCAGTGACATTAGCATTCGGTACTGTTGGTAAATGGAAGACGC 1296
QY 2472 CAAAAACATAAAGAGAGCCGCGCATTCCTCTAGAGGATGGAACCGCTGGAGA 2531
DB 1297 CAAAAACATAAAGAGAGCCGCGCATTCCTCTAGAGGATGGAACCGCTGGAGA 1356
QY 2532 GCCAATCGATTAAGCTTAGAAGATACGCCCTGTTCTCTGGAACAATTCGCTTTACAGA 2591
DB 1357 GCCAATCGATTAAGCTTAGAAGATACGCCCTGTTCTCTGGAACAATTCGCTTTACAGA 1416
QY 2592 TGCACATATCGAGGTGAACATCACGTTCCGCGAATACCTCGAATGTCGGTTCCGTTGG 2651
DB 1417 TGCACATATCGAGGTGAACATCACGTTCCGCGAATACCTCGAATGTCGGTTCCGTTGG 1475
QY 2652 CAGAACTATGAACCATATGGCTGAATACAAATCACAGAATCGTCGATGCAAGTAAA 2711
DB 1476 CAGAACTATGAACCATATGGCTGAATACAAATCACAGAATCGTCGATGCAAGTAAA 1535

Qy	2712	ACTCTCTTTCATTCCTTATGCGCGTGTGGCGCCGTTATTTATCCGAGTTGCAGTTGC	2771
Db	1536	ACTTC-TTCAATTCCTTATGCGCGTGTGGCGCGTATTTAT-CCGAGTTGCAGTTGC	1593
Qy	2772	CGCC	2775
Db	1594	GCCC	1597

Search completed: August 1, 2005, 12:08:58
Job time : 1711 secs

This Page Blank (uspto)